International Journal on Advances in Intelligent Systems

2020 vol. 13 nr. 1&2

The *International Journal on Advances in Intelligent Systems* is Published by IARIA.
ISSN: 1942-2679
journals site: http://www.iariajournals.org
contact: petre@iaria.org

International Journal on Advances in Intelligent Systems
Volume 13, Number 1 & 2, 2020

Suzana Dragicevic, Simon Fraser University- Burnaby, Canada
Mauro Dragone, University College Dublin (UCD), Ireland
Marek J. Druzdzel, University of Pittsburgh, USA
Carlos Duarte, University of Lisbon, Portugal
Raimund K. Ege, Northern Illinois University, USA
Jorge Ejarque, Barcelona Supercomputing Center, Spain
Larbi Esmahi, Athabasca University, Canada
Simon G. Fabri, University of Malta, Malta
Umar Farooq, Amazon.com, USA
Mehdi Farshbaf-Sahih-Sorkhabi, Azad University - Tehran / Fanavaran co., Tehran, Iran
Anna Fensel, Semantic Technology Institute (STI) Innsbruck and FTW Forschungszentrum Telekommunikation Wien, Austria
Stenio Fernandes, Federal University of Pernambuco (CIn/UFPE), Brazil
Oscar Ferrandez Escamez, University of Utah, USA
Agata Filipowska, Poznan University of Economics, Poland
Ziny Flikop, Scientist, USA
Adina Magda Florea, University "Politehnica" of Bucharest, Romania
Francesco Fontanella, University of Cassino and Southern Lazio, Italy
Panagiotis Fotaris, University of Macedonia, Greece
Enrico Francesconi, ITTIG - CNR / Institute of Legal Information Theory and Techniques / Italian National Research Council, Italy
Rita Francese, Università di Salerno - Fisciano, Italy
Bernhard Freudenthaler, Software Competence Center Hagenberg GmbH, Austria
Sören Frey, Daimler TSS GmbH, Germany
Steffen Fries, Siemens AG, Corporate Technology - Munich, Germany
Somchart Fugkeaw, Thai Digital ID Co., Ltd., Thailand
Naoki Fukuta, Shizuoka University, Japan
Mathias Funk, Eindhoven University of Technology, The Netherlands
Adam M. Gadomski**,** Università degli Studi di Roma La Sapienza, Italy
Alex Galis, University College London (UCL), UK
Crescenzio Gallo, Department of Clinical and Experimental Medicine - University of Foggia, Italy
Matjaz Gams, Jozef Stefan Institute-Ljubljana, Slovenia
Raúl García Castro, Universidad Politécnica de Madrid, Spain
Fabio Gasparetti, Roma Tre University - Artificial Intelligence Lab, Italy
Joseph A. Giampapa, Carnegie Mellon University, USA
George Giannakopoulos, NCSR Demokritos, Greece
David Gil, University of Alicante, Spain
Harald Gjermundrod, University of Nicosia, Cyprus
Angelantonio Gnazzo, Telecom Italia - Torino, Italy
Luis Gomes, Universidade Nova Lisboa, Portugal
Nan-Wei Gong, MIT Media Laboratory, USA
Francisco Alejandro Gonzale-Horta, National Institute for Astrophysics, Optics, and Electronics (INAOE), Mexico
Sotirios K. Goudos, Aristotle University of Thessaloniki, Greece
Victor Govindaswamy, Concordia University - Chicago, USA
Gregor Grambow, AristaFlow GmbH, Germany
Fabio Grandi, University of Bologna, Italy
Andrina Granić, University of Split, Croatia
Carmine Gravino, Università degli Studi di Salerno, Italy
Michael Grottke, University of Erlangen-Nuremberg, Germany
Maik Günther, Stadtwerke München GmbH, Germany
Francesco Guerra, University of Modena and Reggio Emilia, Italy
Alessio Gugliotta, Innova SPA, Italy
Richard Gunstone, Bournemouth University, UK

Jeffrey Junfeng Pan, Facebook Inc., USA
Hervé Panetto, University of Lorraine, France
Malgorzata Pankowska, University of Economics, Poland
Harris Papadopoulos, Frederick University, Cyprus
Laura Papaleo, ICT Department - Province of Genoa & University of Genoa, Italy
Agis Papantoniou, National Technical University of Athens, Greece
Thanasis G. Papaioannou, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
Andreas Papasalouros, University of the Aegean, Greece
Eric Paquet, National Research Council / University of Ottawa, Canada
Kunal Patel, Ingenuity Systems, USA
Carlos Pedrinaci, Knowledge Media Institute, The Open University, UK
Yoseba Penya, University of Deusto - DeustoTech (Basque Country), Spain
Cathryn Peoples, Queen Mary University of London, UK
Asier Perallos, University of Deusto, Spain
Christian Percebois, Université Paul Sabatier - IRIT, France
Andrea Perego, European Commission, Joint Research Centre, Italy
Mark Perry, University of Western Ontario/Faculty of Law/ Faculty of Science - London, Canada
Willy Picard, Poznań University of Economics, Poland
Agostino Poggi, Università degli Studi di Parma, Italy
R. Ponnusamy, Madha Engineering College-Anna University, India
Jerzy Prekurat, Canadian Bank Note Co. Ltd., Canada
Didier Puzenat, Université des Antilles et de la Guyane, France
Sita Ramakrishnan, Monash University, Australia
Elmano Ramalho Cavalcanti, Federal University of Campina Grande, Brazil
Juwel Rana, Luleå University of Technology, Sweden
Martin Randles, School of Computing and Mathematical Sciences, Liverpool John Moores University, UK
Christoph Rasche, University of Paderborn, Germany
Ann Reddipogu, ManyWorlds UK Ltd, UK
Ramana Reddy, West Virginia University, USA
René Reiners, Fraunhofer FIT - Sankt Augustin, Germany
Paolo Remagnino, Kingston University - Surrey, UK
Sebastian Rieger, University of Applied Sciences Fulda, Germany
Andreas Riener, Johannes Kepler University Linz, Austria
Ivan Rodero, NSF Center for Autonomic Computing, Rutgers University - Piscataway, USA
Alejandro Rodríguez González, University Carlos III of Madrid, Spain
Paolo Romano, INESC-ID Lisbon, Portugal
Agostinho Rosa, Instituto de Sistemas e Robótica, Portugal
José Rouillard, University of Lille, France
Paweł Różycki, University of Information Technology and Management (UITM) in Rzeszów, Poland
Igor Ruiz-Agundez, DeustoTech, University of Deusto, Spain
Michele Ruta, Politecnico di Bari, Italy
Melike Sah, Trinity College Dublin, Ireland
Francesc Saigi Rubió, Universitat Oberta de Catalunya, Spain
Abdel-Badeeh M. Salem, Ain Shams University, Egypt
Yacine Sam, Université François-Rabelais Tours, France
Ismael Sanz, Universitat Jaume I, Spain
Ricardo Sanz, Universidad Politecnica de Madrid, Spain
Marcello Sarini, Università degli Studi Milano-Bicocca - Milano, Italy
Munehiko Sasajima, I.S.I.R., Osaka University, Japan
Minoru Sasaki, Ibaraki University, Japan
Hiroyuki Sato, University of Tokyo, Japan
Jürgen Sauer, Universität Oldenburg, Germany
Patrick Sayd, CEA List, France

Dominique Scapin, INRIA - Le Chesnay, France
Kenneth Scerri, University of Malta, Malta
Rainer Schmidt, Austrian Institute of Technology, Austria
Bruno Schulze, National Laboratory for Scientific Computing - LNCC, Brazil
Ingo Schwab, University of Applied Sciences Karlsruhe, Germany
Wieland Schwinger, Johannes Kepler University Linz, Austria
Hans-Werner Sehring, Namics AG, Germany
Paulo Jorge Sequeira Gonçalves, Polytechnic Institute of Castelo Branco, Portugal
Kewei Sha, Oklahoma City University, USA
Roman Y. Shtykh, Rakuten, Inc., Japan
Robin JS Sloan, University of Abertay Dundee, UK
Vasco N. G. J. Soares, Instituto de Telecomunicações / University of Beira Interior / Polytechnic Institute of Castelo Branco, Portugal
Don Sofge, Naval Research Laboratory, USA
Christoph Sondermann-Woelke, Universitaet Paderborn, Germany
George Spanoudakis, City University London, UK
Vladimir Stantchev, SRH University Berlin, Germany
Cristian Stanciu, University Politehnica of Bucharest, Romania
Claudius Stern, University of Paderborn, Germany
Mari Carmen Suárez-Figueroa, Universidad Politécnica de Madrid (UPM), Spain
Kåre Synnes, Luleå University of Technology, Sweden
Ryszard Tadeusiewicz, AGH University of Science and Technology, Poland
Yehia Taher, ERISS - Tilburg University, The Netherlands
Yutaka Takahashi, Senshu University, Japan
Dan Tamir, Texas State University, USA
Jinhui Tang, Nanjing University of Science and Technology, P.R. China
Yi Tang, Chinese Academy of Sciences, China
John Terzakis, Intel, USA
Sotirios Terzis, University of Strathclyde, UK
Vagan Terziyan, University of Jyvaskyla, Finland
Lucio Tommaso De Paolis, Department of Innovation Engineering - University of Salento, Italy
Davide Tosi, Università degli Studi dell'Insubria, Italy
Raquel Trillo Lado, University of Zaragoza, Spain
Tuan Anh Trinh, Budapest University of Technology and Economics, Hungary
Simon Tsang, Applied Communication Sciences, USA
Theodore Tsiligiridis, Agricultural University of Athens, Greece
Antonios Tsourdos, Cranfield University, UK
José Valente de Oliveira, University of Algarve, Portugal
Eugen Volk, University of Stuttgart, Germany
Mihaela Vranić, University of Zagreb, Croatia
Chieh-Yih Wan, Intel Labs, Intel Corporation, USA
Jue Wang, Washington University in St. Louis, USA
Shenghui Wang, OCLC Leiden, The Netherlands
Zhonglei Wang, Karlsruhe Institute of Technology (KIT), Germany
Laurent Wendling, University Descartes (Paris 5), France
Maarten Weyn, University of Antwerp, Belgium
Nancy Wiegand, University of Wisconsin-Madison, USA
Alexander Wijesinha, Towson University, USA
Eric B. Wolf, US Geological Survey, Center for Excellence in GIScience, USA
Ouri Wolfson, University of Illinois at Chicago, USA
Yingcai Xiao, The University of Akron, USA
Reuven Yagel, The Jerusalem College of Engineering, Israel
Fan Yang, Nuance Communications, Inc., USA

Zhenzhen Ye, Systems & Technology Group, IBM, US A
Jong P. Yoon, MATH/CIS Dept, Mercy College, USA
Shigang Yue, School of Computer Science, University of Lincoln, UK
Claudia Zapata, Pontificia Universidad Católica del Perú, Peru
Marek Zaremba, University of Quebec, Canada
Filip Zavoral, Charles University Prague, Czech Republic
Yuting Zhao, University of Aberdeen, UK
Hai-Tao Zheng, Graduate School at Shenzhen, Tsinghua University, China
Zibin (Ben) Zheng, Shenzhen Research Institute, The Chinese University of Hong Kong, Hong Kong
Bin Zhou, University of Maryland, Baltimore County, USA
Alfred Zimmermann, Reutlingen University - Faculty of Informatics, Germany
Wolf Zimmermann, Martin-Luther-University Halle-Wittenberg, Germany

International Journal on Advances in Intelligent Systems
Volume 13, Numbers 1 & 2, 2020

CONTENTS

Compassionate Game Design: A Holistic Perspective for a Player-centric Game Design Paradigm for Games4health

Aslihan Tece Bayrak
Dept. of Game Development
Media Design School, MDS
Auckland, New Zealand
tece.bayrak@mediadesignschool.com

Abstract—**Despite the growing interest in using games for positive health outcomes such as alleviating disease symptoms, mental or physical therapies, or rehabilitation, finding suitable games or developing custom games fit for purpose is challenging for the researchers. Using commercial games for these purposes pose various usability challenges and health and safety pitfalls. When it comes to designing one, creating a compelling game is already challenging. This article presents a novel theory called compassionate game design that aims to encourage development of compassionate games with the help of a player-centric game design paradigm. The paradigm facilitates a closer exploration for affordances and limitations of games for players with (mental or physical) health conditions or with diverse physical and cognitive abilities. The main contribution of the paradigm is its ability to promote a closer focus on player resources and encourage a diverse inquiry towards creating games that are more compassionate.**

Keywords - game design; player-centric design; user-centered design; accessibility; games for health; compassionate design.

I. INTRODUCTION

The interest in game based interventions or game assisted treatments for greater health benefits has increased significantly over the last decade with a visible change towards a more positive perception of games. Games for health research hopes to leverage the compelling nature of games for positive health outcomes. For the games aiming to assist medical treatments or provide therapy or rehabilitation, the target player group is special as they are very different than a healthy player demographic. Despite the increase in the variation of player demographic playing all kinds of commercial games [63], the symptoms or the impacts of a disease such as (motor or non-motor symptoms) in Parkinson Disease (PD), or the effects of the conditions (such as mood or attention span in autism) are what makes a player demographic special. This means depending on the complexity of conditions, the ability to perform in a commercial game is limited, and the chance of frustration due to failures during the performance is high [4]. When the main purpose for using a game is increasing adherence to the therapy or treatment, the game needs to be engaging while also fulfilling its core purpose, i.e., therapy, rehabilitation, treatment, etc. [28]. Therefore, games that are designed for this purpose need to be favoring player's conditions while also establishing motivation and engagement. This means, a clear approach towards understanding the needs of the player and usability scope of the game is essential. Usability scope of the game is defined and shaped by its system-like properties that are curated through game design. Therefore, guidance to unite this understanding with game design would be beneficial for successful solutions.

Even though industry practice advices and attempts for early involvement of players by methods such as playtesting, focus group activities, interviews, questionnaires and usability analysis, participative practices in game development are not in the same form or shape as such in product design and development mainly due to the creative nature of intellectual property. Lack of player-centered approaches to game design in commercial contexts was also discussed by other researchers [30][35][54]. In the hindsight, common industry practices and priorities for commercial game development may not be relevant or efficient for developing games for special purposes such as learning or health and rehabilitation. Lack of a suitable game design methodology that explores designing for these demographics seems to be a result of lack of engagement with design research within games studies [21]. Therefore, creating player experiences for special purposes that are embedded in entertainment can seem overwhelmingly complex for a researcher who is not identifying as a game designer. Both for the use of commercial games and in order to develop new games, there is a need for further design research to inform suitability and purpose of the games for health outcomes [4][5][28]. A player-centered model that unfolds the layers of a game experience with respect to the player's perspective can help a designer (or a researcher-designer) cut through this complexity.

This work presents a player centric design paradigm for designing games for health and rehabilitation purposes. The paradigm is player-centric by prompting the designer to contemplate on player perceptions and player capacity. It encourages the designers to reinforce a closer relation across the layers of the game experience while promoting a discussion to understand the relation between players' conditions and the layers of the game. Therefore, it emphasizes compassion for the current state of the player in terms of their disease (especially if they have a chronic disease), mental and physical stamina, or their abilities in order to foster development of suitable game elements for positive health benefits. This means game design is influenced by a player model. In Merriam-Webster, compassion is defined as [a] "sympathetic consciousness of others' distress together with a desire to alleviate it". Thus, the model aims to support creation of a compassionate game.

To present the paradigm, its development and application, the paper is composed of the following sections. Section II presents the motivation and background work to explain why this paradigm is necessary and the plateau of literature to ground the work. Section III discusses the methodology. Section IV explains what compassionate design is and its challenges. Section V presents the player-centric design model with its layers and main elements while Section VI presents a discussion and conclusion.

II. MOTIVATION AND BACKGROUND

This design paradigm emerged during a research through design activity, and further developed into its current form with a synthesis of existing game design approaches and user experience research. A scoping review for player-centric (or player-centered) methods for game design revealed little in terms of an established methodology, techniques or model while the notion of player-centered design is commonly mentioned to be favored as user-centered equivalent of user experience design. Current game design methodologies are unhelpful to contemplate how to facilitate usability at the game level, how each design decision impacts the player's ability to play, and what to focus on during a design activity involving players with health issues [35]. Treadaway et al. [57] initiated a discussion on compassionate design for dementia care and developed guidelines on the use of compassion in activities of design for wellbeing. As much as these guidelines bring important concepts to the spotlight, i.e., dignity and personhood, application of these in game design can be more effective if adopted by a player-centric thinking. Towards synthesizing a player-centric design model for health purposes, this section presents the state of the art on game design, player-centric design, players, player experience, and games for health. The last subsection presents a case analysis to unite the background work towards strengthening the motivation.

A. On game design

Literature has a few well recognized, sometimes loosely defined, frameworks for game design and analysis. These are Schell's elemental tetrad [20], Fullerton's Formal-Dramatic-Dynamic (FDD) elements [19], Adams' interaction model [9] and mechanics-dynamics-aesthetics (MDA) [48]. All of these agree to see games as systems that resemble the characteristics of a closed system with the elements enclosed, their properties, and relations between those elements [13][66]. Among these, Fullerton's formal elements are a core set of elements granular enough to allow a micro-level thinking for the design or analysis of a game [19]. Formal elements allow studying the game grammar by dissecting a game into its smallest building blocks. These are listed in Table 1. with a brief explanation on how they relate to game grammar. Notice that all formal elements are somewhat interrelated in their definition, role and use in game design.

In Fullerton's FDD model [19], dramatic elements bring meaning to the formal elements and establish context while dynamic elements emerge when formal elements are put into motion.

TABLE 1. FORMAL ELEMENTS OF GAMES AND THEIR DESCRIPTION; REPHRASED FROM GAME DESIGN WORKSHOP BY FULLERTON [19].

Formal elements of games	*Description*	*Supporting questions*
players	The main agents of the game; define interrelations among players, the effects of rules on players, etc.	How many players, what are their roles, and how do they interact with one another?
procedures	Methods of play and actions towards achieving the game objectives	What exactly the players do, how do they do that?
resources	In-game assets that can be used to accomplish goals; they are a part of in-game economy; they must have utility and scarcity.	What can I afford? How is a resource useful to the player? When and how the player can access/lose resources?
objectives	Represent the purpose of the player within the rules of the game.	What are the players trying to achieve? What kind of impact do the objectives have on the game?
rules	Define game objects, actions that are allowed and restricted; rules govern the system.	What is the player allowed to do? How are the rules presented and enforced?
conflict	Emerges from the oscillation between goals and rules (accomplishing the goals within the given rules is a challenge).	What is the challenge? What makes it hard/challenging to accomplish the goals?
boundaries	Define the physical and virtual play-space by separating game elements from anything that does not belong to the game.	What are some physical limitations that define the play space? What are some conceptual limitations?
outcome	The result of the activity/session/game; this must be visible and measurable.	How does the game end? How does the player progress/advance? Is the objective reached?

However, the relation between those are rather loosely defined or are not presented in a form that would easily allow their use by un-trained game designers. Refraining from a detailed discussion on the set of formal elements, both the elemental tetrad and the MDA framework present a high-level model based on visibility of the layers or different perspective holders. Elemental tetrad and derivatives of this model [36] show a four-way relationship between mechanics, aesthetics, story and technology drawing attention to how visible these are from outside the system. Even though this four-way relationship is helpful for game analysis, it is too loosely defined to be helpful as a game design approach. On the other hand, MDA presents two separate viewpoints; one from designers' perspective, the other from players'. This makes MDA a useful model to discuss different perspectives even though it is not helpful for the design process. For a game designer, mechanics are designed to give rise to dynamics, and hence aesthetics are conveyed as a result of these dynamics. Therefore, they see mechanics, dynamic and aesthetics in this given order. Here, aesthetics represents all emotional aspects that stem from the game. Conversely, player's perception follows an inverse order: aesthetics-dynamics-mechanics. It is framed with their perception of the aesthetics; therefore, their viewpoint is always under the influence of their perception of the aesthetics. As much as elemental tetrad and MDA are helpful for contemplation on perspectives, they do not facilitate a closer investigation towards dismantling a complete game experience into its thinner slices, explaining influences for experience design or designing a game for special purposes and/or special demographics such as people with chronic health issues. Such granularity is missing and would be very helpful for designing compassionate games that would keep emphasis on the capabilities of the player.

Almeida and Silva [41] reviewed a collection of game design methods and tools (as per 2013) drawing attention to the use of visual languages and design modelling despite favored varyingly across industry and academic fields. They argue that the immaturity of design research in game design field prevents the adoption of these by researchers or industry. Moreover, they note that the game design methods and tools under review lack representation of the player perspective, especially for aesthetic components. Indeed, these approaches seem to mostly favor the game over the player, therefore lacking methods to draw attention to the player and facilitate compassionate design. For this, a visual game design model can be helpful for communication across researchers, designers and other members of the development team while developing compassionate games.

B. *Player-centric Design*

It is commonly thought that the player-centric design is an extension of the user-centered design. On one hand, one can argue that game design is inherently player-centric since it always questions what the player is doing, what they are allowed to do, their objectives, rewards, the consequences for the player, and how the player is supposed to feel during the gameplay experience. Therefore, it would be unfair to suggest that game design is not player-centric. On the other hand, game design practice is a creative endeavor, meaning that it may carry a separate authorial intent due to its creative nature. In the virtual boundary of play-space, the game world is governed by the rules and dressed in the metaphors of the game. Besides, game design practice could also be perceived as play-centric, putting emphasis on gameplay over other elements of the game. Therefore, rather than the habits, requests or expectations of the player, the nature of play and how it manifests itself may take precedence. In response to all these arguments, player-centric design is centered with empathy to the player and aims to provide a positive experience to the player despite contradictions with creative pursuit [9]. Therefore, player-centric design puts the player before the creator.

Charles et al. also [27] acknowledge that most of the player-centered approaches to game design rely on the research on human factors and are adopted from user-centered approaches. Although these are useful for addressing usability issues at an application level, the additional qualities of games as a media (such as the level of challenge, player skills and gameplay, etc.) are not addressed. Adams presented another model that is based on the interaction between the player and the game, yet does not encourage explorations on sensory complexity of games more than perceiving them as user interfaces [9]. Although this could be seen as a much clearer approach that allows the use of existing HCI research for games, it creates further ambiguity on the potential needs of games on a visual and contextual level.

When game design literature is scanned for the role of player in game design, it becomes apparent that player-centric game design is not as common, but involvement of players in design process such as in participatory design is even less considered or troublesome [30][54]. A more often seen application of player-centric design is with early playtests and the use of player personas rather than participatory design sessions as seen in product design or service design. As much as iterative design practice is praised, most of the time studios prefer to involve as few people as possible and as late as possible due to secretive nature of creative production. Sometimes an audience model is created to enable intellectual assumptions for players' needs assuming that designers have a decent understanding of what players wants and needs. This notion has been strongly criticized for neglecting the real persona of the player and reducing them to simple unimportant actors disconnected from a larger socio-cultural context. Sotamaa suggests designers to focus on what they are good at while acknowledging the players as the specialists of their own "everyday gaming" [35].

On the other hand, player involvement in context of player-centeredness can also be considered continuous rather than finalized by the end of production period of a game. Player modelling [27] and automatic/dynamic difficulty adjustment based on biosensors [53] are some methods to create more player focus during gameplay. Industry practice ackowledges the use of personas rather than focus groups [66] since the involvement of individual tastes of potential players may be noisy rather than useful for the designer even though persona development or the personas themselves are rarely shared by the game designers within the industry [44].

Inspired from Norman's emotional design [12], Baharom et al. [33] suggest a conceptual framework that offers a valuable discussion about the relation between emotion, perception, reason and structural elements of games. The framework is encouraging for discussions on the cognitive, psychomotor and affective dimensions of games by placing player-centric design in the center although interrelations between these dimensions remain unclear. Hodent also argues that user experience (UX) should not be separated from the design cycle as it is the glue that ensures intended experience and players' perceptions work together to create an enjoyable gameplay experience [56].

C. *Unpacking Play and Players*

The concept of play and why people play have been a hot topic for decades in social and cultural studies, psychology, and philosophy. Sutton-Smith argued for the cultural significance of play stating that it takes many forms throughout the life of each individual [65], and Suits argued that games are central to human existence and that "everyone alive was really a grasshopper in disguise" [64]. Game studies as a relatively young research field also inherits this question, and further asks why people enjoy playing a game more than another, what it is that people find in the activity of gameplay, what motivates them for this activity, how this enjoyment is governed, how it can be shaped, etc.

1) Aesthetics of play: Aesthetics, as adopted from MDA framework by [48], represents emotional aspects communicated by the game. Previous section on game design already mentioned that emotions emerge via the dynamics of the game experience, and mechanics of a game serve for the aesthetic goals. The aesthetics mentioned by [48] are (broadly) sensation, fantasy, challenge, narrative, fellowship, discovery, expression, and submission. On the other hand, Lazzaro argues that we play because it is fun despite the ambiguity in the definition of "fun". She suggested the Four Keys to Fun method to explain the phenomena of "fun" [49]. In 4Keys2Fun, fiero (hard fun), curiosity (easy fun), excitement (serious fun), and amusement (people fun) are chosen as four core emotions. Lazzaro presents groups of emotions as subset of these four and discusses how some actions (towards creating the game features) can resolve into these emotions. For example, actions of collection, rhythm and repetition create emotions of excitement, focus and relaxation. Therefore, they create "serious fun". The suggestion is building a play experience to evoke more than one emotion to prevent stagnation.

2) Meaningful play: Salen and Zimmerman [13] state that "the goal of successful game design is the creation of meaningful play." Meaningful play emerges from the relation between a player's action and the system outcome that is visible as immediate and long-term effect. If this relation is not clearly visible to the player, players lose interest in the gameplay. If the relation is discernable immediately and the result is integrated into the larger context of the game, the actions in the game become meaningful. With immediate feedback, the player receives a confirmation about their action; with long term effect, player sees that their choices matter and their actions carry meaning towards reaching their in-game goals. Hence, meaningful play motivates the player experience.

3) Motivation: Psychology studies on human behavior consider that factors that activate and sustain behavior towards a goal or that create propensity to learn and act are motivational factors, and are either intrinsic or extrinsic in nature. Intrinsic motivation comes from within such as enjoyment in the activity itself rather than for an external reason whereas extrinsic motivation originates from an external reward [45]. If extrinsic motivation is considered as controlled motivation, through practice this can be transformed into self-determined autonomous motivation (hence intrinsic). Self Determination Theory (SDT) suggests that human motivation is based on three basic psychological needs: competence, autonomy, and relatedness [37]. Competence refers to the universal will to perform better at something; autonomy refers to the desire to be in control of one's own path; relatedness refers to the desire to connect with other beings. Although SDT originates from the psychology field, Player Experience Needs Satisfaction (PENS) model [38] translates the dimensions of SDT into measurable domains for game development and analysis. According to PENS, when these three needs are fulfilled for a player within the context of the game world, players experience higher levels of satisfaction, and potentially continue playing the game.

4) Player personalities and preferences: The idea of player types and that some features are favoured by some types of players were first mentioned with the first Multi User Dungeon (MUD) [47]. Player types (sometimes referred as Bartle's player types)—killers, achievers, socializers, and explorers were representing players' interest in MUDs. Killers are interested in features that allow them to act on other players such as combat and direct interaction; achievers are interested in features that resolve with tangible acknowledgement of success such as badges, unique items, etc.; socializers are interested in features that encourage playing with other players, collaboration and social interaction; explorers are interested in adventure, discovering and mostly acting on the world. This is later adopted by other designers and researchers and expanded upon [51][52], still somewhat staying true to the first four types but developing interim types to suit the variety of game experiences present today.

As much as continuation of play is of interest, it is argued that there is a taste factor involved in a player's initial choice to start playing [39]. Correlations between motivation and taste may be possible although not supported by any research. Regardless, VandenBerghe argues that taste factor can be obtained by taste maps as presented in Engines of Play method [39].

Engines of Play is a method suggested by VandenBerghe [40], initially at a game developers conference. It relies on the Big Five psychological model (also known as Five Factor Model or OCEAN) and is suggested as a way to find out a correlation between game features and personality of players to enable a "more accurate empathy" towards different kinds of players. There is no empirical validation reported by the date of this writing; nonetheless, the idea is that people seek out experiences that align with their personality traits. Five personality traits are mapped to five discrete domains (OCEAN): openness, conscientiousness, extraversion, agreeableness, and neuroticism. This approach was further developed into taste maps to understand player motivations (domains of play) in relation to personality traits and game features by VandenBerghe [40]. Later, in his blog he noted that the Engines of Play method was employed by another game studio to analyze their existing franchise as an early exercise prior to adopting it in the future, and that the method was successful in terms of clarifying the reasons behind the success of the franchise and allowing all team members to effectively communicate intent and result at various stages of development by using the stereotypes created with the taste maps [55].

Stereotypes mostly take the form of personas and inform the design process while creating contextual layer and aesthetics. In a post about player profiles for *Magic: The Gathering* (Wizard of the Coast, 1993), Mark Rosewaters (head designer of *Magic: The Gathering*) explains that personas (he refers to them as psychographic profiles [44]) allow him to isolate different personality traits and behaviors so that what motivates a particular type of person to act in a certain way could become clear. He continues "[to] create an emotional response, I had to understand what emotions I was trying to evoke". Therefore, personas can be used to guide the design strategy with empathy so that suitable gameplay features can be developed.

D. Unpacking Game Experience

Game experience research shows interest in understanding the emotional experience of a player so that a better and more purposeful player experience can be created [62]. Towards this goal, user-centered design encourages the use of focus groups, participatory design, usability testing, and usability heuristics; however, the reception of these methods in games industry is mixed [59]. From a commercial angle, many studios seem to consider some of these methods less effective or time/resource consuming. It is hard to argue against business oriented decisions; however, this notion seems to be less effective when developing games for health and learning. Commercially, a common thought in game development is accepting that the product will not be appealing to all users within the target demographic, therefore should be selectively designed for the most profitable subset demographic. Nonetheless, in the context of games for health and learning, this notion can be challenged for the purpose of maximizing the reach of the product. Therefore, a closer look at the concepts of game experience is necessary. These are optimal experience, usability, playability, accessibility, and game feel.

1) Optimal experience: A balance between skill and challenge is often argued to create an optimal experience for a person in any kind of performance; flow [50]. Challenge depends on the fine balance between an individual's ability to perform up to the standard of a given objective. When they do not have the competence to fulfil the demands of the activity, they feel anxiety. Flow theory explains that there are seven core components to flow. They are categorized under conditions and characteristics. Conditions (clear tasks, feedback, concentration, attainable goal) are the prerequisites of flow while characteristics refer to the experience of the individual in the state of flow. They are control, diminished awareness of self, altered sense of time. According to Csikszentmihalyi, maintaining this balance is one of the most critical prerequisites for flow, and it is important for both motivation and learning [50]. In order to maintain engagement and promote flow, challenge should be matching player growth in skill, tasks and goals should be made clear, feedback should be timely and readable, and distractions should be minimized.

2) Usability: Usability in game development aligns with common usability guidelines developed for human computer interaction even though usability for games and game-like experiences encompass more than just the interface or the interaction modalities. Nielsen and Norman group defines usability as "a quality attribute that assesses how easy user interfaces are to use" [6]. A usable interface should be easy for a user to become familiar with and competent in; to achieve their objective; to discover previously less known attributes when need arises; to recover from an error; and to recall how to use the interface on subsequent visits. As much as these can be valid for a game interface, it is necessary to note here that player experience (game experience) is larger than the interface itself; therefore, an isolated perspective that only focuses on the experience based on interaction modality and the attributes of user interface would highly likely miss the bigger picture for evaluating the experience of a game. The bigger picture includes all facets of player experience including but not limited to interaction modalities, user interface, audio-visual design, in-game challenges and objectives, player attention, engagement, immersion, etc. As Juul and Norton [46] point out, distinguishing the difficulty in interaction/interface and the difficulty of gameplay as a challenge of the game is necessary to separate game usability from that of production software. Since games are by nature inefficient and lusory, the line between difficulty in interface versus the challenge of a game could be rather blurry while a gameplay challenge can also be located at the interface.

3) Playability: Extending from usability research, playability and player-game interaction has been gaining some speed to explore practical applications for game usability challenges [16][26][34][62], and there is further work on playability and the complexity of player involvement in player experience evaluation [25]. Sanchez et al. [25] define playability as "a set of properties that describe PX [Player eXperience]", and add that playability measures the effectiveness, efficiency, satisfaction and fun of performing to a specified performance as per the specific goals of the entertainment experience. Even though the same attributes of usability are applicable to playability, their meaning are different. Since the main goal of PX is entertainment, Sanchez et al. collated seven attributes to characterize playability: satisfaction, learnability, effectiveness, immersion, motivation, emotion, and socialisation. Among these, motivation and emotion are directly about the user and shaped by player preference. In addition, motivation is also said to be "a key factor in generating a positive experience for the player.

4) Accessibility: Despite the well-defined nature of accessibility within user experience domain, accessibility may differ for special demographics [33]. Accessibility of a game refers to the ability to enable the participation of people with disabilities to interact with the game and play without feeling any barriers during their interaction with the game or during overcoming the challenges presented by the game. Isbister states that "video games should be considered above all as complex IT systems, with which individuals wish to interact within the framework of a goal-directed activity" [34]. In fact, the complexity of each differs depending on pre-designed layers of gameplay experience, which may be open to emergence of more complex behaviours in some cases. On the other hand, accessibility seems to be simplified to interaction modalities without much consideration on cognitive load, emotional demand or context related issues. There is no disagreement on the importance of a suitable interaction modality and a satisfying user experience, and that they are integral to the creation of engaging and enjoyable game experiences. Furthermore, it is essential for a player to be able to see past the interaction modalities while playing a game so that they could fully engage with the activities and events in the game, hence be in the flow of the game. Juul emphasizes that challenge of the gameplay should never be mixed with interface difficulty [46].

5) Game feel: Game feel, in Swink's definition, is "the tactile, kinesthetic sense of manipulating a virtual object" [10, pg.xiii], and it is essential to the game experience. Three building block of game feel—real-time control, simulated space, and polish, shapes the game feel towards resolving into five common experiences. They are (as taken from [10, pg.10]):

- The aesthetic sensation of control;
- The pleasure of learning;
- Practicing and mastering a skill;
- Extension of the senses;
- Extension of identity;
- Interaction with a unique physical reality within the game.

The pleasure of the game experience mostly maps to a combination of those. For example, with driving games, it is more about the aesthetic sensation of control; yet, while driving in an obstacle course, it is more about practicing and mastering a skill. When the car is damaged, and the player screams at that moment, it is about extension of the senses and identity. This is also how game feel connects to SDT. As a player gets better at performing in the game, their competence increases. As their competence increases, they will feel and demand more autonomy for their actions. The combination of balanced competence and autonomy (resembling optimal experience) makes them motivated. As they feel motivated and focused on the experience, they will feel connected to their avatar or their car that they are driving; hence, their avatar becomes a part of their identity.

E. Games for Health and Rehabilitation

Over the years, literature on games for health has grown with both the use of commercial games and the development of custom made games for health benefits including but not limited to improved balance and gait, muscle strength, reach, etc. [2]-[5]. Games are considered as a solution to non-adherence to therapy, the lack of motivation for therapy and/or rehabilitation, the service and delivery costs of therapy such as unbalanced patient-therapist ratio, reach to and distribution of services, lack of accuracy in tracking progress, etc. There are convincing arguments [28] about using commercial games for physical therapy since they are already products that are commercially developed to create and maintain engagement; however, clear contributions on re-appropriation of existing games for specific purposes and maturity of guidelines to adapt those for people with chronic diseases are limited. Similarly, researchers have mostly adopted or applied existing game design models [9][19], some extended from existing models [18][31], or developed on an ad-hoc basis in order to create a suitable product without much exploration on the design practice [4][5][14][15]. Lack of a design methodology to facilitate design exploration for a selected demographic with unusual limitations, such as motor, cognitive or mental difficulties that may be developed due to a chronic disease, seems to be a result of limited design research within games research [21]. This means, learnings and observations from these studies are fragmented and hard to reconcile for further studies. Both for the use of existing commercial games and with the purpose of developing custom games, there is a need for further design research to ensure suitability and accessibility of games for special demographics [4][5].

Potential recipients of game-based rehabilitation include a variety of age groups across a variety of diseases such as cerebral palsy, ADHD, autism, post-stroke, Parkinson's and parkinsonism, age-related deterioration, etc. Development initiatives for games for health in each of these seem to be discrete and isolated. Even though the requirements of each and every rehabilitation and limitations of each and every target user group are different, it is not unrealistic to expect a common design strategy or a framework that can accommodate common methods to develop games that augment rehabilitation during the treatment of these diseases.

Within games for health context, a set of guidelines was collated to inform developers on the special constraints for their design [11][14][17][18] while also pointing out the lack of further development in the game design discipline. Game design in its practice is agile and opportunistic, yet the design practice itself is under-researched. A study by Isbister and Mueller on variety of interaction modalities especially in the field of movement based games is one of the few that evaluates conditions of interaction and suggests strategies for a more successful design of movement-based systems, mainly games [17]. The guidelines provide insights for the design of interaction modalities, necessary feedback compatible with the interaction device and activity, extends from self-expression, challenge and fun. Among the few studies that attempt to develop games for PD patients, Assad et al. previously suggested design principles for designing games for PD, particularly for motion based games for rehabilitation purposes [11]. Although informative, the principles are prescriptive for a specific type of game (such as arm extension based exercise features or balance features) rather than allowing a wider applicability supported with a clear methodology for game design. Moreover, while each item in the guideline can be affecting various areas in the game, the interconnected nature of the game experience as per these affects is hard to perceive. Therefore, a model such as the one presented in this paper could have been useful to guide the designer towards realising these relations. For example, fostering wider and fluent movement is advised, but input modality is not discussed probably because the guidelines are aimed for movement based games. Gerling et al.'s "extended model" (seem to be extended from Fullerton's formal elements [19]) presents a useful basis for designing for older adults [18]; however, the analysis lacks a wider perspective on the interplay between the elements of the model in relation to a player-centric focus. As a game design paradigm, such perspective would encourage an exploration of player's perspective in this context especially for the needs/issues of special demographics with health issues. Nevertheless, the notes on the player's abilities (both cognitive and motor abilities, such as attention span, short-term memory, repetitive input, etc.) to be considered as resources resonates with the perspective of compassionate game design concept presented here. In line with the previous discussion here, Burke et al. [29] suggested three principles to game design for health and rehabilitation: meaningful play, handling failure and setting a suitable level of challenge, and applied these to the development of games for the rehabilitation of stroke patients. Despite limited reflection on the development process or the design iterations, the note on adaptive aspects of the games and that they were well-received by the players is valuable.

F. *Case analysis: On Design Challenges for Players with Impairments due to Parkinson Disease*

In light of state of the art presented above, this subsection presents a case analysis to unite the background work towards strengthening the motivation. In this example, the case analysis looks into the difficulties that people with PD may be facing in their daily life. In the context of compassionate design, these difficulties help addressing the design challenges. A similar analysis would be useful, in fact necessary, while attempting to design for a special demographic. Later, in the presentation of the paradigm, the notes from this analysis is used to show how challenges are addressed within the holistic nature of the paradigm.

Research shows that quality of life for PD patients drops over time meaning that activities of daily living, such as dressing, grooming, bathing, self-feeding and functional mobility are jeopardized as disease develops [23]. Even at earlier stages various disturbances and impairments limit the ability of the users while performing tasks that are considered simple; rendering many games inaccessible for this player base. The disturbances and impairments that are commonly observed across PD patients are sensory sensitivities, motor impairments, cognitive impairments and emotional sensitivities. Many of these impairments, especially in early stages of PD, show close similarities to age related changes that are commonly observed among older adults or other health related situations, such as stroke patients [14].

1) Sensory sensitivities: Sensory difficulties include not only hearing or vision problems as mostly seen with elders [18], but also sensitivities for sensory overload due to visual and sound stimuli. Occupational therapy for PD advise reducing visual stimuli by reducing confusing patterns (striped-checker), strong colours and hues, and simplifying the load by preventing contrasting visuals and clutter [22]. Visio-spatial disturbances and strong contrasts cause freezing while clutter overloads cognitive processes with a need of strategising and replanning. There are no specific sensitivities reported about audio; however, the use of metronome and inducing percievable rhythm into daily life are presented to be useful to enhance motor abilities [22].

2) Motor impairments: Main motor impairments observed in PD are trembling fingers and hands (tremor), rigidity, slowness in movement (bradykinesia), and gait problems [22]. Subtle slowness in movement, postural change and gait problems are also seen in elders even though the scale of these differ from PD. Trembling fingers and hands, especially depending on the scale of movement can make it very hard to use an input device or perform button presses while the slowness in movement can increase the response time.

3) Cognitive impairments: Cognitive impairments that are commonly seen with PD are learning and retaining information (working memory), concentration and attention, and executive functions. Executive functions are a set of inter-related cognitive processes that are essential for goal-directed behaviours [24]. Even though they are heavily related to cognitive domains, motor skills and connection between cognitive and motor skills are the main reason why they are absolutely necessary for activities of daily living [24][23]. In order to preserve gait, a person needs to evaluate their surroundings, strategically decide a path of movement, shift their weight and meanwhile check their balance. If they come across an obstacle, they should be able to stop executing their plan and rework a new plan similar to the correction cycle mentioned before. This means all six executive functions are actively used during a simple walking task: attention, inhibition, planning, reasoning, shifting (flexibility), and working memory. Gait disorders share similar issues originating from deficiencies in executive functions and also observed among older adults. This means impairments in executive functions also develop among older adults, perhaps milder than PD. In addition to a previously identified need for task creation frameworks to facilitate purposeful use of games for special demographics [4], it has become appearant that there is a need for a design paradigm that draws attention to the abilities and limitations of the players, thereby clearly mapping the position of the player in relation to the game experience. Therefore, the player-centric design model is developed to support and inform designing games for rehabilitative and preventative therapies for PD.

III. METHOD

The method of this study is design science research (DSR) while the presented work aligns closer with the idea of developing a nascent theory through the means of DSR [42]. Despite the emergence of new frameworks or design principles towards facilitating purposeful use of games for serious means such as creating positive change, alleviating disease symptoms, chronic disease rehabilitation, etc. [28], there is a clear fragmentation in the design field, separating other more flourished design fields and game design. The compassionate design philosophy and the player-centric game design paradigm are developed through grounding strategies [42], and the synthesis draws from human computer interaction, user-centered design practices, and game design. As per the activities of DSR compiled by [43], the activities of this research are below.

- *Activity 1*: Problem identification and motivation is discussed in Motivation and Background section (Section II) while conceptual grounding is also presented. Existing approaches for the intersection of game development, user-centered design and the state of the art for design approaches for games for health are presented. Core issues of research projects utilizing games for therapy/rehabilitation of special demographics including chronic diseases and elders are (1) researchers' potential lack of knowledge and understanding for game design and development towards creating a compelling game experience or selecting a suitable game, (2) the complexity of a game experience that would fulfill the requirements of a therapy (or rehabilitation), and (3) the ability to identify the necessities and expectations of the chosen demographic within the context of game design. Hence, the conceptual grounding not only provides a motivation for the problem but also synthesizes theory from the material sources in order to build grounding material for the suggested paradigm.

- *Activity 2*: In relation to the problems identified in the previous section, objectives for a solution are discussed in the Compassionate Game Design section (Section IV). The core objective is drawing attention to three areas about the player: (1) player-game relation, (2) player resources, (3) player perception and game content. The objectives of the compassionate design are explained starting from these three areas while value grounding for these are also enclosed. Following from the core problems identified with the activity 1, the activity 2 reinforces the development of the compassionate game design concept. Hence, a leading philosophy for the design paradigm is developed via an inquiry on purpose and action.
- *Activity 3*: In response to the objectives identified during activity 2, the player-centric design paradigm is developed to present an easy to apply model for development and selection of games for a special purpose. The paradigm presents a clear account for the layers of a game experience, how these layers are interconnected, and the role of the player in relation to these. The details of the concepts across the model, their nature and their relations are explained in the Section V.
- *Activity 4*: The purpose of the paradigm is enabling the development or selection of a game fit for purpose. This activity is dispersed across the explanation of the model in Section V. A detailed discussion for the application of the model as per the holistic capacity of the paradigm is presented. The application of the model, design examples and validation fall out of the scope of this paper and are planned for future work.

The work stands as a theory building DRS activity to create synthesis across diverse fields while providing an easy to follow perspective for researchers, who may be new to games for health or who may have limited knowledge about game design. The artefacts of the activities are the compassionate game design concept (presented in Section IV) and the player-centric game design paradigm (presented in Section V). Strengths, limitations and future steps are discussed in the last section.

IV. COMPASSIONATE GAME DESIGN

What is meant with *compassionate game* may seem fluffy and idealistic. However, the idea of a compassionate game originates from the comfortable and embracing feeling of compassion. Compassion has been actively practiced within the context of healthcare and made its way into public healthcare policies despite the ambiguity of its application as compassionate care [61]. In healthcare context, the meaning of compassion sometimes gets mixed with dignity, identity, and self-worth. Regardless, it is suggested that caring conversations and building relationships with the patients can help to effectively practice compassionate care [61]. If we are considering games as platforms to deliver or support healthcare, we need to start practicing ways of integrating compassion into game design.

A compassionate game would tolerate the failures of a player with kindness rather than confrontation so that the failures are perceived as opportunities of development rather than unsuccessful attempts; therefore, the remaining feeling is not frustration. A compassionate game would allow for a grace time that suits a player's capabilities rather than confronting the player with a challenge that they cannot improve upon; therefore, the remaining feeling is not a decline in self-esteem. Compassionate games would not compete with the player on a layer that the players cannot improve; they rather encourage the player for participation by providing a harmonious and suggestive environment of support and care. Therefore, they strive for accessibility, inclusivity, care, tolerance and adaptation. Compassion in this sense may seem to conflict with the notion of challenge and conflict in game design. On the contrary, it is about finding the right area in the game design where challenge should be planted so that it is possible for the players to fairly overcome the challenge rather than facing an unsolvable/impassable problem due to their limitations; thus, feel incompetent. Therefore, compassion does not suggest removing the challenge but planting it in an area where player progression is not limited by the disease. Hence, it is about care and tolerance with an understanding of what a player is going through so that their competence develops.

Applying compassionate game design means seeking further inquiry in the design process in relation to three areas about the player: (1) player-game relation (at a one-to-one level), (2) player resources (skills and abilities), (3) player perception and game content (what is going on in player's mind versus what the game is offering). By doing so, the player model, as the central element of compassionate game design will be integrated into the process of developing a game that is compassionate towards its players. The following subsections present the compassionate design concept as per these three areas while the last subsection discusses the challenges of compassionate design.

A. Player-game Relation: Interactivity

Crawford [66] defines interactivity as "a cyclical process in which two actors alternately listen, think, and speak. The quality of interaction depends on the quality of each of the subtasks (listening, thinking, and speaking)." This points to the relational prospect of interactivity and its iterative nature.

The *micro level* (the smallest interactive cycle, i.e., moment-to-moment) and the *macro level* interactions (a larger trajectory of interactions) are intertwined to create a meaningful player experience [13]. Within compassionate design context, a compassionate game would show competence in listening more attentively, finesse in speaking with care, and less complexity in the combination of interactions and interaction modalities. Therefore, the design paradigm needs to draw attention to the flow of interaction.

Citing from Sutton-Smith, Salen and Zimmermann [13] refer to game experience as a combination of five dimensions: visual scanning, auditory discriminations, motor responses, concentration and perceptual patterns of learning. Players scan the entire scene based on the visual and auditory signs while concentrating on events and signals provided by the game. They perform actions based on the demands of the game and proceed whilst scanning for visual and auditory cues. This cycle continues as the player carries on playing the game. As they do so, players learn more about the patterns of play and improve their understanding of the game from familiarity to higher expertise. At its core, the cycle of play is an implicit micro game loop that stands on the perceptual understanding of the game world, the processing time of the perceptions, and the response from the player. Norman suggests that the ideal interaction cycle is present when no psychological effort is necessary to use a system (bridge the gulf of execution) [58]. This resembles the correct game feel notion in which the performance of interaction with the system becomes a compelling spatial presence [10].

Swink explains the interaction cycle as an action-correction cycle where player perceives an interim goal towards completing an action and slowly reaches their target, except this happens so fast that we do not necessarily notice the process [10]. The trouble arises for people with neurodegenerative diseases. Due to motor or/and cognitive impairments, a part or some parts of the action-correction cycle show delays; therefore, the action becomes harder to perform. When there is delay at system's end, we see that as a usability problem due to the lack of timely feedback. When delay happens at player's end (Fig. 1), system needs to be gracefully waiting and in fact supporting at times by additional supportive feedback. Since our main concern is reducing frustration and increasing compassion in our design, we need to consider this cycle both for the player and the game.

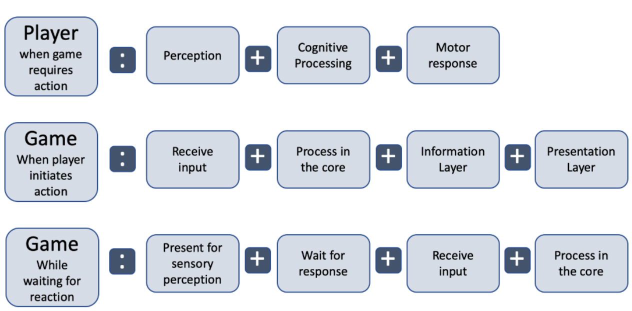


Figure 1. Interaction between player and the game when an action is initiated by a player or a game

The figure above (Fig. 1) presents the stages of the interaction between player and the game (action correction cycles from top to bottom):

- (1) When it is player's turn, what happens at player's end: Player perceives the situation, it goes through their cognitive process, then player creates a motor response; meanwhile, the game waits.
- (2) When it is game's turn after player initiates an action, what happens at game's end: Game receives input, processes the information in its structural core (games as a closed formal system), then contextualizes the outcome with information layer, then feeds the information to the presentation layer via relevant metaphors.
- (3) When game responds for player's turn while waiting for a reaction from the player, what happens at game's end while waiting for player's reaction: Game presents the current state of the game and waits for an input (response of the player); then receives the input and passes it to its core to process as response again. From the moment an input is received, the rest of the cycle is the same as the second cycle. When the game reaches to the second step here—waiting for response, this step is where the first action cycle takes place.

Each of these interaction cycles requires well-tuned graceful delay for player action (the moments that the game receives input), and player perception (the moments that the game presents output). Intended player actions or player reflexes should be considered to be potentially affected by their condition, so should player perception—as it includes vision, hearing, and cognitive processing that may also be affected by their condition. This means response from the game should be perceivable by the player with longer and lingering feedback to support the moment-to-moment gameplay within an action-correction cycle and ensure a continuity of perception. More about the response of the game and forms of the response are discussed as the paradigm explained in Section V (Subsection D).

B. *Player Resources (Skills and Abilities): Game Difficulty, Challenge and Accomplishment*

Starting questions for identifying the right challenge for a compassionate game experience goes through identifying the areas impacted by a disease or condition. A thorough analysis for the underpinnings of impairments, disease symptoms, physical or emotional difficulties enhances the start of a better-informed game design process. This is exemplified in the case analysis presented before. Previous discussion on formal elements (Section II.A) already mentioned the holistic nature of these elements. While the conflict between the in-game game objectives and the rules form the challenges in the game, difficulty sits in a sliding scale based on what kind of skills are necessary and the complexity of the procedure to perform in-game actions. Player skills, abilities, and capacity shape the player performance, therefore player experience. This means the paradigm needs to draw attention to these for a true player-centered approach.

Skill, as per the definition we use from the dictionary[1], is "dexterity or coordination especially in the execution of learned physical tasks"; ability[2] is "the quality or state of being able"; and a resource loosely refers to the capacity. In this context, the players suffering from a condition such as motor impairments may have the skill (originally) to perform fine motor tasks such as buttoning a shirt; however, their ability may not be allowing to do so any more. This is most of the time true for all patients as they go through different stages of a chronic disease [60]. For diseases such as PD, Alzheimer's, Dementia, etc., cognitive skills, motor skills, emotional and physical stamina deplete over time as they progress.

Player resources pose a larger problem depending on patient demographics. For example, Parkinson's patients develop cognitive impairments such as working-memory issues, or reduced flexibility and inhibition, which may hinder their strategic planning ability. Even when they are able to perform a level of strategic planning, how fast they get tired in the process of re-strategizing restricts their capacity in this task, therefore their resource for strategic planning. They also develop motor-skill impairments such as precision control, bradykinesia, freezing, etc. As disease develops, the scale of these hindrances grows demanding more support and compassion from the system. Since these impairments limit their cognitive and motor ability to perform precise, repetitive or complex tasks, this means their cognitive and motor resources to perform in this system is limited. In addition, their physical resources such as endurance, agility and stamina are limited along with their emotional resources such as resilience, self-efficacy, and self-esteem. Hence, it is important to identify the resource intensity of the skills (type and scale of resources—motor, cognitive, emotional, physical) necessary to perform to a level of completion.

1) Motor resources: These refer to the capacity in performing activities with fine and/or motor skills such as quick-response, timed-response, precision, repetition, reach, etc. If the ability of a player is already restricted in any of these domains, interaction modality needs to accommodate the restriction. For example, people with tremor may not be able to use a control scheme that requires timed-response and precision simultaneously. What is and is not achievable by the player needs to be identified in the process.

2) Cognitive resources: These refer to the capacity in performing activities that demand cognitive abilities such as remembering a number of items for a certain amount of time (working memory, short-term and long-term memory), attention (selective attention, sustaining and retargetting attention), anticipation (pattern recognition and prediction), decision making, problem solving, planning (sequencing, prioritization, reordering), etc. People with conditions that affect cognitive resources, executive functions are also affected [24]; hence, quality of life is under threat. For example, people with attention deficit hyperactivivity disorder (ADHD) will find it hard to sustain their attention; therefore, a game designed for this target demographic needs to reclaim their attention repetitively and be ready to gently remind without drawing attention to the reality that they did not pay attention in the frst place.

3) Emotional resources: In playability context, Sanchez et al. [25] refer to the emotion as players' involuntary response. In the context of compassionate game design, emotional resources refer to the capacity in perseverance while feeling those emotions evoked by the game. It is less likely that a player would find it hard to deal with too much fun; however, it is more likely that they will have a limit for frustration before they give up playing the game. Thus, the emotional resources are resilience, self-esteem, self-efficacy, trust, etc. The capacity of the emotional resources can be influenced if not directly affected by a disease or a condition. The response of a game at moments of success and failure may drain or refill this resource.

4) Physical resources: These refer to the capacity in physical performance including the ability to perform throughout the gameplay session (endurance), ability in combined tasks such as movement and button press (dexterity), speed, muscular strength, power, etc. This may sometimes overlap with motor resources especially when precision and reach are required as in interactive dexterity by Sanchez et al. [25].

Overall, failing forward and a positive reinforcement throughout the challenge as per the previous discussions on motivation and optimal experience, should be the driving force of a compassionate game. According to dopaminergic studies [28], it is noted that uncertainty of result (win or lose) and almost succeeding or narrowly avoiding failure increase not only motivation but also dopamine release.

C. Player Perception and Game Content: Motivations and Reservations

Playing a game is rewarding; it rewards the player with an experience of "fun" as long as they wanted to play it. Swink suggests a challenge for a game designer is to create a game mechanic that is worth learning for players [10]. In fact, the difficulty of learning a procedure for gameplay only becomes worthwhile by the return of that action and how integral that action for the continuum of gameplay. Therefore, rewards of mastery, how difficulty of a procedure scales for the variety of individuals with diverse skills and abilities, and the relation of the mechanic to the contextual layer of the game (its art, theme and story) shapes the motivation. Meanwhile, pre-judgement, previous experience or fear shapes reservations. If the reservation grows larger than what motivations has to offer, it causes non-adherence. For a compassionate game, the challenge is creating pathways to overcome the reservations. This goes through positively changing the player perceptions. Games have the unique ability to invite people in and persuade them for participation. Their ability to shape perception and fight reservations of the participants should be explored further towards creating compassionate games.

D. On Challenges of Compasionate Game Design

Within the context of compassionate game design, a game that is not able to accommodate a player's health or learning

[1] https://www.merriam-webster.com/dictionary/skill

[2] https://www.merriam-webster.com/dictionary/ability

related difficulties is considered to be impaired. Therefore, the main challenge of compassionate game design is removing the impairment of the game so that it is playable by special demographics. This starts with identifying areas that players are struggling due to their conditions. As a starting point, these areas can be categorized in three domains: sensory challenges, physical challenges and cognitive challenges. For example, a core challenge of game design for special demographics with health issues is the restrictions on the physical and cognitive abilities of the player group. When this merges with the perceptions of the players about the nature of the activities, it becomes hard to create motivation to play the game. Therefore, it is important to notice that when players are not able to fulfil the expectations of the moment-to-moment gameplay simply because it is not possible due to their current condition, they will lose interest in the game as the game is only another reminder of their inability; hence, the experience is not enjoyable. Then, the challenge is introducing tolerance and support for these moments of need.

According to Swink [10], any delay that breaks the continuity of an experience creates poor game feel. Therefore, a game needs to be responsive for the player inputs. Similarly, any player action in response to the events presented by the game needs to be timely; otherwise, deemed unsuccessful by the game. Hence, the game acknowledges successful behavior and rewards it while also clearly communicating the consequences upon failure. From a purely game design point of view, facing consequences helps to reinforce meaning in choices. However, from a player-centric design point of view, especially for players with motor and/or cognitive impairments as in Parkinson's or stroke patients, how much time should be evaluated as the time-frame for "timely response" is unclear. The procedures of the game should be forgiving with a suitable error margin and compensating for delays as motor-cognitive processes in player's mind may take longer than an ordinary player. This may mean shifting the challenge to another aspect of game rather than using input-timeliness, input-combinations, or input-related challenges as a central element of core gameplay. Especially for anxiety prone demographics, removing time-constraint related design elements such as time-trials (obvious time constraint), falling platforms (hidden time constraint), traversal challenges (time constraint imposed by combination of other elements such as enemies, moving platforms, etc.) all together could even be a more suitable approach. In consideration of mental barriers, such compassion would reduce anxiety and stress, while replacing those with confidence, self-assurance and trust.

Game accessibility requests compassion from the system to bridge the gap between the player and the game in order to reduce the impact of impairment. For example, similar to how a poor visual design of an interface would reduce the usability significantly, a poorly developed game world with low quality art assets and poor choice of color palette would also reduce the quality of game experience. For a user interface (UI), chosen color palette, negative space, density/scarcity of visual elements, placement and alignment of UI elements such as button, text boxes etc., consistency of all UI elements in and across the pages including font, size and color, and the feeling and readability of transitions are extremely important. All these elements come together to create the right feeling for the interface with its content and suitability to the context. Similar but more intense, contextual content of games requires far more depth compared to a user interface. The player perceptions of the game world not only stem from the narrative elements that are telling a story of the game world but also the familiarity of the player with the game and game world. Therefore, the player needs a high processing power in order to evaluate all the information they could gather from the game in real time while playing the game. Thus, complex visual stimuli may become paralyzing if the system does not show enough compassion when a player's impairment is getting in the way of their play activity. Without a model that helps breaking down areas that demand processing power and areas to hide delay, it is uneasy to contemplate on how this need could be resolved without frustration.

V. A Holistic Model for Player-centric Game Design

The model seen in Fig. 2 attempts to draw attention to the identified core areas of compassionate design as per the discussion in Section IV. It aims to encourage a designer to take a player's perspective in relation to the formal structure of games and aims to enable a deeper discussion on game design challenges for special demographics. In doing so, the purpose is presenting a holistic map laying down the building blocks of the compassionate game design concept. This concept becomes readable with a game model and a player model that interacts with one another. With a player-centric focus, the model (Fig. 2) is extended from [1], and combines layers of player experience—as inspired from Garret on the planes of user experience [7][8]—with Adam's game design model [9] and Fullerton's formal elements [19]. In addition, Mayra's [67] core and shell dialectic informed the work for explaining the relation between formal elements of a game and a game's context. In a simplistic fashion, the dialectic captures the actors of player experience to start a conversation about their relation even though it lacks details to explain the potential elements of player experience or player's role in this dialectic.

The rationale for compassionate design is discussed in the previous section including a perspective for potential challenges it may entail. Here, player-centric game design model is explained with its layers.

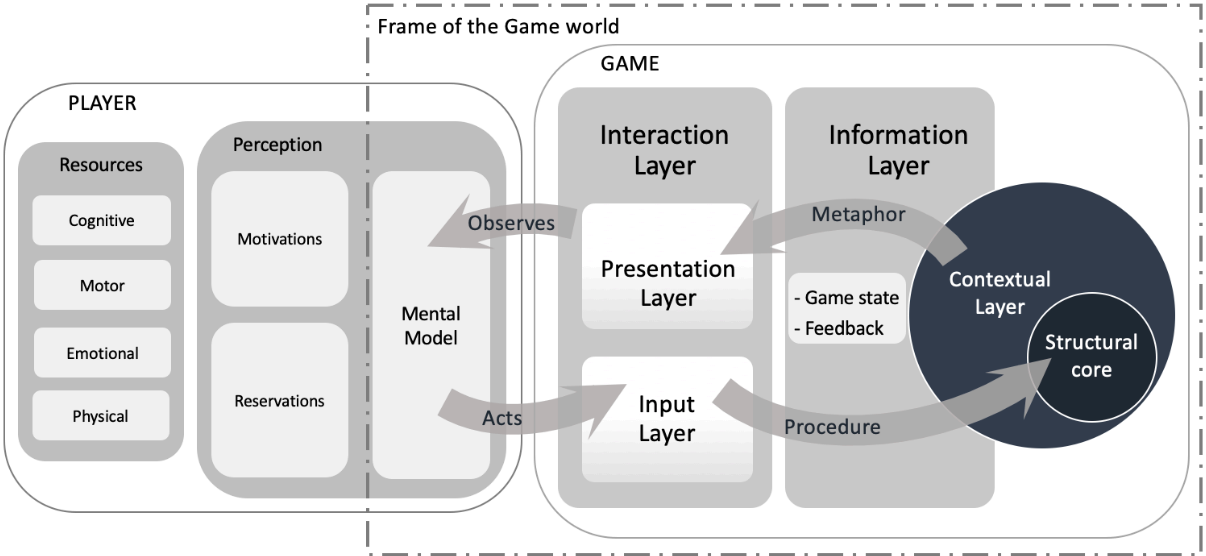


Figure 2. The player-centric game design model aims to encourage the development of compassionate games (extended from [1]). Therefore, it favours a player model by bringing transparency to the capabilities of the player and helps a designer to perceive the system from players' point of view.

With this model, a designer (i.e., researcher as a designer) can map the systematic flow of player experience and the mechanisms carried out in the game without losing focus on the player end of the game world. When the interaction cycle starts with a presentation of a game scene to the player, player observes it through their perception. As they act via pressing the buttons on the controller, moving the joystick or somewhat invoking the input layer as per the interaction modality available to them, the game receives this input and carries it to the structural core of the game to be resolved as per the predefined rules (as per the formal elements of the game). When this is resolved, the response will be dressed up by the contextual layer (in order to fit with the theme and style of the game) and shaped into the form of necessary information as per the game, its genre and style; hence, a responsive feedback will be passed to the presentation layer for the player to observe and start the interaction cycle again. Compassion is a necessity to be practiced at all stages of this cycle in order for a game to become compassionate.

A. *Frame of the Game World*

Frame of the game world acknowledges that the player model overlaps with the game model in creation of the gameplay experience similar to the discussions on experience of play in Salen and Zimmermann's schemas [13]. Schemas explain how the experiential layer of a game emerges when players play it. The game is governed by the rules that are formal elements of the game grammar and stays same for each player, yet the experience of each player is unique as per their session filtered by their perception. Each player engages with the game world through their perceptions that are mainly based on their mental model [12]. In this context, the relation between the game and the player is initiated and maintained by the player. Therefore, the frame of the game world encloses the player as it is activated with the presence and interaction of the player.

B. *Player Model*

Player model is composed of player perception and player resources. It considers the player perception as a combination of motivations, reservations and a mental model through which players engage with the game world. Swink refers the player perception as "perceptual field" [10] that is a combination of all previous experience including ideas, thoughts, memories, etc. From a player's point of view, game is only the interface since they do not need to know about (or perhaps are not interested in) any inner workings of the system (Fig. 3).

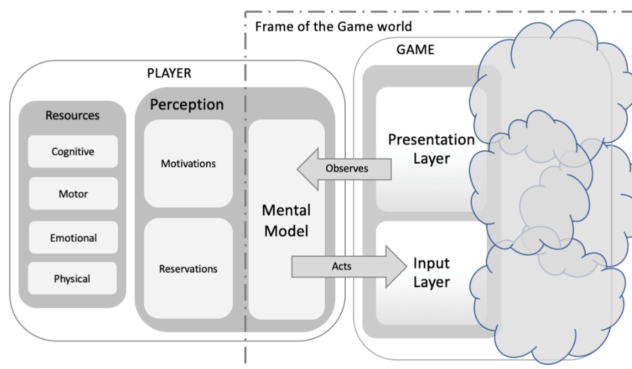


Figure 3. When we truly think in a player-centered fashion, it becomes apparent that players are very likely not aware of the interconnected system that makes the game function.

Game presents the current state of the system to the player, this is observed and filtered through the mental model to be judged against motivations and reservations by player's perception. Once there is a decision or a resolution at player's end, an action is performed by the player via the input scheme of the game. The game receives this input and magically processes it. Both decision making process and performing the action are taxing on player's resources.

1) Mental model: Mental models refer to thought processes about how something works in the real world. They are formed via observations, or assumptions based on prior knowledge. Each person's mental model is clouded with their pre-judgements and their interpretation of the system [58].

There are three models of the system (as adopted from [58]): (1) how the player thinks the game works; (2) how the game is presented to the player as; (3) what the game really is. Mental model maps to the first model, even before player starts any interaction with the system. This resembles the user image and the system image concepts of Norman [58]. After the first interaction, mental model is reinforced with their continuous interaction. The further the gap between the first model and the last—what player thinks for how the game works versus what it really is and how it works—, the lower is playability—hence the higher is frustration.

A well developed game interface either resembles an already known system to leverage familarity or helps the players to develop a clear mental model of what they can do in this game. Tutorials are generally used as a bandate when game is not able to intuitively help players to build a good mental model of the system. The ability of the game to suggest a clear mental model reduces the complexity of interaction during stages of the action cycle [12]. Unlike HCI's common approach of mainly relying on user interface for evaluation of usability, for games human action cycle is dispersed over gameplay. Therefore, mental model is influenced by three contextual elements at perception level: (1) the contextualisation of actions at goal formation stage of the human action cycle, (2) the procedures and their contextual presentation at execution stage of the human action cycle, and (2) audio-visual feedback and game state presentation within user interface (or heads-up display—HUD) at evaluation stage of human action cycle. These are directly related to the presentation aspect of the interaction layer and indirectly related to the information layer as the contextual metaphor is fed from the core of the game.

2) Motivations: Elaborating on the aforementioned research on player motivation, the purpose of this in the framework is to draw designer's (researacher's) attention to the motivations of the player and how they influence their participation. This should align with not only the purpose of the game but also the intentions of the player to motivate participation. With meaning and purpose built in the game, the idea is leveraging the players' motivation for working towards their in-game goals to fulfill the health and rehabilitation goals. This is why *PokemonGo* (Niantic, 2016) has been praised by health practicioners. For a patient who needs to do a descent amount of walking every day, the in-game goal of hunting pokemons serves the purpose of the therapy by motivating the patient to walk since hunting requires finding the pokemons in the first place. In order to find them, a player needs to physically walk around until pokemons appaear in the location based augmented world. Hence, daily walking activity is fulfilled with the help of the mobile game.

3) Reservations: Any reason that prevents a player from participation is referred within reservations. It may be based on previous experience or can be formed at the first instance of interaction. Barriers to physical activity such as fear of injury, poor self-image, low self-esteem or lack of confidence to join group exercise activities are commonly reported among reasons for lack of motivation for therapies or rehabilitation [28][60][63]. Another set of barriers may also originate from unfamiliarity with technology, lack of clarity and ease of use (accessibility) of a system, perceptions on the usefulness of technology, self-image, and fear of failure. Therefore, the player model draws attention to the relation between reservations and motivations, and how they may influence the mental model of a target user group. As much as motivations act as a driving force, reservations act as the hand brake. Nevertheless, the game experience needs to be fluid and effortless.

C. Player Resources

In this context, player resources are considered as the currency that a player spends during the play session. Each individual has finite physical, emotional and cognitive capacity. Player resources represent this capacity and its availability during gameplay. Rather than the resources commonly discussed as a formal element of a game, the resources referred in this model are intrinsic to a player. Intrinsic player resources are cognitive skills, motor skills, physical abilities and emotional abilities (physical and emotional stamina). These are within the power of the player and do not belong to the in-game economy. They are not generated by the game or in the game; however, they are brought in and used by the player, yet consumed by the game. They are limited, and refresh time for these vary from person to person. For a suitable design targeting players with physical or mental health related difficulties, designers need to study how these resources are effected by the disease and how they will be used in the game in order to prevent frustration during the game.

For example, after long hours of play, the physical stamina of a player may drop, and they may not be able to function as prompt as they were at the beginning of the play session. This is an example of a player spending their physical stamina (their physical ability resource); hence, getting tired after a long period of playing. If the player is suffering from anxiety and the game is designed to alleviate this, the decision-making process in the game needs to accommodate that and should not rely on dilemmas as a challenge, dynamics such as time trials or quick-time events should be avoided, and visual stimuli should be organised to prevent clutter and overloading.

D. Interaction Layer

Interaction layer represents the overlapping space of game world and player's world. Via this layer, player sends input to the game world, observes the results of their input and receives a response from the game world. Interaction takes place between the player and the game world. It is either started by the player via an input, or by the game world via an event presented with the presentation layer. When it is initiated by the player, the game responds; when it is initiated by the game via presentation layer, the player responds via input layer. In both cases, information layer feeds the presentation layer with necessary data. Interaction layer is composed of presentation layer and input layer.

1) Presentation Layer: Presentation layer can also be thought as sensory layer encompassing audio, visual and haptic presentation of the current status of the game. This includes continuous presentation of the game world and the game's response to the player inputs. Any feedback generated in response to the player inputs is presented by the presentation layer. Clarity of presentation and a suitable composition of audio-visual elements are essential for readability. The presentation of this layer is either (1) active or (2) reactive:

a) Active presentation (1) is initiated and conciously generated by the game with respect to or regardless of player positions or existance. It is important to notice that in any moment that the player is not interacting with the game, this presentation would still be actively present; hence, active presentation. It is an audio-visual presentation for the theme and of the game world, and would include any audio-visual element to inform the player about the current active or passive status of the game including but not limited to telegraphing for potential actions/events or signalling for awareness of moment-to-moment development of events in the game, score, etc.

b) Reactive presentation (2) is developed in response to or because of an action initiated by a player, therefore can be considered as system's (game's) response to the player; commonly referred as feedback. Necessary to note that this presentation would only be executed if an action that would trigger a reaction was performed by the player. Reactive presentation can take forms of audio-visual and/or haptic feedback directly in response to a player input or to draw attention to a change that was a result of the player action/input. Audio-visual effects in user interface that inform the player for the moment of change are also examples of this category. For example, a falling platform would only fall (or destructed into pieces with an appealing vfx and sfx) if the player character walks/jumps onto it. Therefore, audio-visual presentation (vfx and sfx for distruction) is a reaction to player's interaction with a game element (falling platform in this case). If player health changes during this action, an audio-visual effect on user interface (one of the hearts blinking prior to disappearing) would also be used to draw attention to this change in addition to the update of player health on screen. Any of the reactionary feedback is time sensitive and loses its value at the end of the perceivable window (i.e., 240ms for a full correction cycle). Swink [10] suggests, if a computer's response takes equal or longer than half of the correction cycle (i.e., 120ms), the systems feels laggy.

The response of the game world could be:

- in the form of visual effects that shows the moment and result of interaction (mostly referred as visual feedback);
- in the form of sound that acknowledges the moment and result of interaction (mostly referred as audio feedback);
- haptic such as vibration of the controller;
- in UI including sound and visual effects.

The response fulfills the expectations of gulf of evaluation. The longer it takes the game to respond or the more disconnected the response is to the current perceivable context of the game, the larger is the gulf of evaluation. Hence, the game usability is lower.

2) Input Layer: Input layer is responsible for the interaction device, input techniques, clarity of input mapping, directness, sensitivity and consistency of input. A common goal for a game controller is effortless use with which the input device feels like the extension of the body [10]. This motto becomes challenging for special demographics with impariments or difficulties related to motor skills. For example, PD has some symptoms, such as tremor, bradykinesia or hypokinesia that may cause difficulty in using an input device or perform an input action within a required time-frame or precision. Therefore, additional research into input devices and interaction modalities would be useful. Moreover, further research on familiarity of the player group, mental model, and restrictions of disease stages for input modality is expected to improve the design.

In addition to the input device or interaction paradigm, the complexity of input is also important. The complexity of input refers to the combination of buttons for an input and the characteristics of the combination such as time-specific input, simultaneous or ordered button presses, etc.

To reconcile, an effective interaction layer for a compassionate game needs to be: 1) immediate; therefore, presentation layer needs to present the feedback immediately, and show the change in game-state if there is; 2) responsive; therefore, the presentation layer always presents a response to any input and the player is never left in dark in terms of what is happening with the system; 3) informative; the presentation is readable and clearly presenting all necessary information such as game-state or contextual clues, feedback is clear and adequate; 4) forgiving; the input scheme allows for a feasible input window, error recovery has necessary feedback and is motivational.

E. Information Layer

Information layer sits in the middle of presentation layer and structural core of the game, and interprets outputs of the core system in a readable format for the player. Cues for meaning making (semantics), affordances and limitations for the player [12], contextual visual material, data organization, response of the game (feedback for the player), and any information, such as score, status, outcomes, etc., belongs to this layer. Even though the information is generated by the structural core of the game, its interpretation is handled in information layer and passed to the presentation layer for the player to see. Therefore, the collaboration between information layer and presentation layer carries high importance.

1) Contextual Layer: Contextual content works with the formal elements and supports meaning making. Objectives, rules, procedures, affordances and limitations become meaningful with the help of contextual content. While being important for engaging the audience, contextual content is also important for maintaining attention and motivation. In order to develop content based on the interests of the target audience (elders, kids, young adults, etc.), a participatory approach or persona studies would be preferable. Salen and Zimmermann [13] emphasize that without context, there is no meaning; hence, contextual meaning helps to develop meaningful play. For example, *Pokemon Go* creates a context around training a special creature called pokemons and battling to defend your dojo; therefore, one needs to capture the pokemons prior to training them. This means the activities of walking around to find the pokemons, performing the moves, and using the means to capture them serve within this context. The rarity of the creatures, what kind of environment they prefer, and how they respond the moves of the player feed into the lore of the game. Swink suggests that the visual content and the context need to work together with the help of the right metaphor for a fuilfilling experience [10]. The metaphor represents the connection between context and mental model.

2) Game State: Game state presents any necessary trackable data that influences players' decision making while they are playing. Game state includes but not limited to score, item count, health, current or remaining time, active or planned tasks/quests, current progress, success/failure, win/lose, etc. Transparency of game state is necessary to ensure competence and autonomy in gameplay.

3) Feedback: Feedback refers to a compassionate reaction that is both contextually suitable and positively perceivable by the player. The positive nature of feedback is emphasized in games for health circles [28][11]. Feedback needs to be focusing on the successes rather than failures. Encouraging continuation of participation and reinforcing flow are within the role of feedback. Necessary to note that feedback can take many forms within a game including but not limited to audio effects, audio-visual changes to the entities in the game (sound, color changes, color flashes, outlining, scaling, etc.), animation of entities in the game or animation at user interface level, visual effects, narration, or any change in the game world, etc. Some important things to consider for the attributes of feedback are; how often to show, how amplified it needs to be, how long it should be shown, and in which form (audio and visual) it needs to be. These attributes should also take into account the capability of the player (as per the player resources dsicussed before) for a compassionate feedback. For example, for a player who can be disturbed by the nature of visual stimuli (i.e., colour, shape, light or change in those), visual feeback should be carefully catered not to breach the acceptable visual qualities.

F. Structural Core

Structural core of a game comprises formal elements [19]. Both the information layer and the presentation layer are dependent on the core structure of the game while also creating meaning for it. From player's point of view, the core of the game may be completely invisible (Fig. 3) as their perception is shaped by the presented information (based on how it is interpreted by the information layer). Therefore, discoverability, learnability and consistency of the system need to be resolved at this layer so that relevant data could be fed to the information layer.

1) Procedures: Procedures are the first point of interaction with the input from the player. They are integral to moment-to-moment gameplay, and they define the chain of moves necessary for performance in the game. Reiterating the previous discussion on correction cycle, a delay in any stage of player performance will make the time-frame of user input longer. For example, double jump could require hitting the jump button twice within a second in order to perform double jump. This seemingly simple action could be unexpectedly challenging for a person with rigidity or sloweness of movement, who might find hitting the same button twice hard to repeat within a second's window. Another example is the number of steps necessary to do something, such as the steps to be performed to bake a cake. For a person with memory issues such as in Alzheimers, remembering those would be really hard, therefore frustrating to perform. Thus, procedures should be catered for a suitable grace time, simple recovery (this does not mean game needs to be easy), shorter and less complex chain of actions for ability to learn and retain information. On top of these, additional consideration for impairments would make a big difference. When done right, the core game demands less player resources and/or compensates for lack of those when necessary.

2) Resources: Management of in-game resources may prove challenging if learning, managing and controlling the resource features of the game are challenging for the players' resources as per discussion in Section IV.B. Additionally, depending on the way in-game resources are used in the game, an interface complexity may also occur creating additional sensory and cognitive challenges at a level of presentation layer. For example, games of real-time strategy genre are more resource heavy than single screen puzzle games even though their cognitive complexity may resemble one another.

3) Objectives: The role of objectives in game design has already been discussed throughout the article. An additional note here is on how in-game objectives of the game need to somewhat serve the targeted health and rehabilitation benefits of the overall experience. Previously discussed motivation example on *Pokemon Go* shows how in-game objectives motivate gameplay, and partially serve for the therapy objectives (walking daily). If the therapy objective was doing some moves in addition to walking, the game would have needed these moves performed towards catching a pokemon. Besides, objectives should be discernable and achievable by the player considering limited cognitive and emotional resources may make an objective unreachable.

4) Rules: Since the system is governed by the rules, transparency and consistency of these were already mentioned as necessity. It is important to note that if the complexity of the rules prevent the players from understanding the moment-to-moment gameplay, game fails to be compassionate. The limitations emposed by the cognitive and emotional resources also shape the player's ability to understand the rules.

5) Conflict: Conflict in context of compassionate design requires additional care. By definition, conflict represents the challenge that attempts to prevent a player from reaching their in-game goals and objectives. This is also what separates digital games and game usability from other interactive digital experiences. While challenging a player, a game presents the player with a problem to be solved and to be overcome. Referring back to motivation and the experience of flow, the design of the challenge directly affects the playability of the game. Therefore, developing a suitable challenge and scaling the difficulty for sensitive demographics is absolutely necessary to motivate play and induce flow. Designing the challenge is essentially designing the game.

VI. Discussion and Conclusion

This section shares a discussion and a closing note on the theory and application of the model presented in the paper.

A. On Compassionate Game Design

The idea of compassionate game design emerged during the analysis of potentially suitable commercial games for rehabilitation of PD, followed with an attempt to identify the sources of the issues found with them [4]. During the prototyping stage of an exercise game for PD, the building blocks of the player-centric design paradigm surfaced as a nascent theory. Compassionate game design concept is novel and brave although naïve. The concept has a novel perspective to carry the idea of empathy towards a more applicable format that is much clear for inquiry and contemplation for game design process. It draws attention to the player's position in the game experience, and encourages the researcher/designer to examine the design in relation to three areas about the player. These are player perception, player resources, and player-game relation (discussed in Section IV in detail). Player-centric game design paradigm is the artefact of the design research activity that is grounded with the compassionate game design concept. The model aims to be easy to read and apply by anyone without prior game design knowledge. As any novel idea, the way to see whether it is useful for the larger community is by putting it into use and letting it evolve with the findings.

B. Strengths, Limitations, and Future Work

The purpose of the model is to promote further discussion on the elements of game design with a player-centric focus; therefore, the main strength of the work originates from the incorporation of the user experience model and "game feel" to ensure this. The paradigm welcomes exploration in those layers, and encourages analytical thinking towards a player-centric design. Currently, there is no other work that brings a player model and a game perspective together as a game design model. The details of the player model with the never-ending dance between motivations, reservations, and mental model contributes to the understanding of what players think when they start playing a game. Furthermore, the resources in the player model allows the researcher/designer to keep an eye on the internal challenges only apparent to the player that are of a sensitive nature affected by a condition or a disease.

As much as the compassionate game design concept and the paradigm were developed with a synthesis of current work, the player-centric design model should not be considered to carry any predictive abilities for a best possible system of game based rehabilitation. The development of a compassionate game with the help of the model do not ensure an effective rehabilitation, nor a single design that is useful for all kinds of rehabilitation. A planned improvement is expanding the model with the addition of a set of questions on what to do for each layer, and exemplar design snippets to further help the researchers. Since the bulk of the theory is presented through synthesis now, there is room to develop a user map or a guide for holistic thinking in future work. In addition, a report for the application of the paradigm with a development journal, the resulting product and a postmortem is planned for the future. Finally, despite some mentions of playability, an additional angle with a more detailed playability concept following from existing playability literature, and a discussion on how playability relates to this model are in the pipeline of future work.

C. Conclusion

In this paper, a player-centric design paradigm is presented to improve game design practice for health-related purposes. The paradigm is holistic as it draws attention to the interconnected nature of a game experience, the layers of the game experience to allow a closer understanding of players' perception of these layers, and some approaches to enhance empathy for player resources. The paper encloses an analysis of the current literature for similar game design pursuits, presents a discourse for compassionate game design, and explains the player-centric design paradigm in detail with grounding strategies. The main contributions of the work are the compassionate game design concept and the player-centric design model. They are developed to provide an easy to follow perspective for researchers, who may be new to games for health or who may have limited knowledge about game design despite a vision on using games for health and rehabilitation purposes. The hope is more effectively leveraging the potential of games for this vision with the creation of more compassionate games.

References

[1] A. T. Bayrak, "A Player-Centric Game Design Paradigm with Compassion," *CENTRIC 2018, The Eleventh International Conference on Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services*, pp. 58-63, 2018.

[2] D. Levac, S. Glegg, H. Colquhoun, P. Miller, and F. Noubary, "Virtual reality and active videogame-based practice, learning needs, and preferences: A cross-Canada survey of physical therapists and occupational therapists," *Games for health journal* 6, no. 4, pp. 217-228, 2017.

[3] A. E. Staiano and R. Flynn, "Therapeutic uses of active videogames: a systematic review," *Games for health journal,* vol. 3.6, pp. 351-365, 2014.

[4] A. T. Bayrak, B. Wünsche, and C. Lutteroth, "Reviewing the Evidence: In Pursuit of a Framework for Parkinson Disease Rehabilitation with Games," *The Promise of New Technologies in an Age of New Health Challenges*, p. 9, 2016.

[5] B. Bonnechère, *Serious Games in Physical Rehabilitation*, Springer, 2018.

[6] J. Nielsen, "Nielsen norman group," [Online]. Available from *https://www. nngroup.com*, last accessed May 2020.

[7] C. Macklin and J. Sharp, *Games, Design and Play: A detailed approach to iterative game design*, Addison-Wesley Professional, 2016.

[8] J.J. Garrett, *Elements of user experience, the user-centered design for the web and beyond*, Pearson Education, 2010.

[9] E. Adams, *Fundamentals of game design*. Pearson Education, 2014.

[10] S. Swink, *Game feel: a game designer's guide to virtual sensation*. CRC Press, 2008.

[11] A. Oliver, R. Hermann, D. Lilla, B. Mellies, R. Meyer, L. Shevach, S. Siegel et al. "Motion-based games for Parkinson's disease patients." In *International Conference on Entertainment Computing*, pp. 47-58. Springer, Berlin, Heidelberg, 2011.

[12] D. A. Norman, *Emotional design: Why we love (or hate) everyday things*, Basic Civitas Books, 2004.

[13] K. Salen Tekinbaş and Eric Zimmerman, *Rules of play: Game design fundamentals*, MIT press, 2004.

[14] M. Balaam, S. R. Egglestone, G. Fitzpatrick, T. Rodden, A. Hughes, A. Wilkinson, T. Nind et al. "Motivating mobility: designing for lived motivation in stroke rehabilitation." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3073-3082. 2011.

[15] A. C. de Barros, J. Cevada, À. Bayés, S. Alcaine, and B. Mestre, "User-centred design of a mobile self-management solution for Parkinson's disease," In *Proceedings of the 12th international conference on mobile and ubiquitous multimedia*, ACM, p. 23, 2013.

[16] L. Caroux, K. Isbister, L. Le Bigot, and N. Vibert, "Player–video game interaction: A systematic review of current concepts," *Computers in Human Behavior* 48, pp. 366-381, 2015.

[17] K. Isbister and F. "Floyd" Mueller. "Guidelines for the design of movement-based games and their relevance to HCI," *Human–Computer Interaction* 30.3-4, pp. 366-399, 2015.

[18] K. M. Gerling, F. P. Schulte, J. Smeddinck, and M. Masuch, "Game design for older adults: effects of age-related changes on structural elements of digital games," In *International Conference on Entertainment Computing*, Springer, Berlin, Heidelberg, pp. 235-242, 2012.

[19] T. Fullerton, *Game design workshop: a playcentric approach to creating innovative games*, CRC press, 2008.

[20] J. Schell, *The Art of Game Design: A book of lenses*. AK Peters/CRC Press, 2014.

[21] A. Kultima, "Game design research," Proceedings of the 19th International Academic Mindtrek Conference, ACM, 2015.

[22] A. Aragon, B. Ramaswamy, and J. C. Ferguson, *The professional's guide to Parkinson's disease*, Parkinson's Disease Society, 2007.

[23] G-M. Hariz and L. Forsgren, "Activities of daily living and quality of life in persons with newly diagnosed Parkinson's disease according to subtype of disease, and in comparison to healthy controls," *Acta Neurologica Scandinavica* 123.1, pp. 20-27, 2011.

[24] A. Diamond, Executive functions, *Annual review of psychology,* vol. 64, pp. 135-168, 2013.

[25] J. L. G. Sánchez, F. L. G. Vela, F. M. Simarro, and N. P. Zea, "Playability: analysing user experience in video games." *Behaviour & Information Technology* 31, no. 10: 1033-1054, 2012.

[26] J. L. Sánchez, N.O. Zea, and F.L. Gutiérrez, "From usability to playability: Introduction to player-centred video game development process," In *International Conference on Human Centered Design* (pp. 65-74). Springer, Berlin, Heidelberg, 2009.

[27] D. Charles, A. Kerr, M. McNeill, M. McAlister, M. Black, J. Kcklich, A. Moore, and K. Stringer, "Player-centred game design: Player modelling and adaptive digital games," In *Proceedings of the digital games research conference*, vol. 285, p. 00100. 2005.

[28] K. Lohse, N. Shirzad, A. Verster, N. Hodges, and H. M. Van der Loos, "Video games and rehabilitation: using design principles to enhance engagement in physical therapy," *Journal of Neurologic Physical Therapy*, *37*(4), 166-175, 2013.

[29] J. W. Burke, M. D. J. McNeill, D. K. Charles, P. J. Morrow, J. H. Crosbie, and S. M. McDonough, "Designing engaging, playable games for rehabilitation." In *Proceedings of the 8th International Conference on Disability, Virtual Reality & Associated Technologies*, pp. 195-201. 2010.

[30] O. Sotamaa, L. Ermi, A. Jäppinen, L. Laukkanen, F. Mäyrä, and J. Nummela, "The role of players in game design: a methodological perspective," In *Proceedings of the 6th DAC Conference*, pp. 34-42. Copenhagen: IT University of Copenhagen, 2005.

[31] B. De Schutter, "Gerontoludic design: extending the mda framework to facilitate meaningful play for older adults," *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS)* 9, no. 1, 45-60, 2017.

[32] K. Miesenberger, R. Ossmann, D. Archambault, G. Searle, and A. Holzinger, "More than just a game: accessibility in computer games," In *Symposium of the Austrian HCI and Usability Engineering Group*, Springer, Berlin, Heidelberg, pp. 247-260, 2008.

[33] S. N. Baharom, W. H. Tan, and M. Z. Idris, "Emotional design for games: A framework for player-centric approach in the game design process," *International Journal of Multimedia and Ubiquitous Engineering*, *9*(10), 387-398, 2014.

[34] K. Isbister and N. Schaffer, *Game usability: Advancing the player experience*. CRC press, 2008.

[35] O. Sotamaa, "Perceptions of Player in Game Design Literature," In *DiGRA Conference*, 2007.

[36] J. G. Bond, "Introduction to Game Design, Prototyping, and Development: From Concept to Playable Game with Unity and C," Addison-Wesley Professional, 2014.

[37] R. M. Ryan, C. S. Rigby, and A. Przybylski, "The motivational pull of video games: A self-determination theory approach, " *Motivation and emotion*, *30*(4), 344-360, 2006.

[38] S. Rigby and R. Ryan, "The player experience of need satisfaction (PENS) model," *Immersyve Inc*, 1-22, 2007.

[39] J. VandenBerghe, "Engines of play," In Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play-CHI PLAY, Vol. 14, pp. 476-476, 2014.

[40] J. VandenBerghe, "Engines of Play: How Player Motivation Changes Over Time," In *GDC Vault*. [VIDEO]. Available from: https://www.youtube.com/watch?v=Lg2GndSat1E&feature=youtu.be, 2016, accessed May 2020.

[41] M. S. O. Almeida and F. S. C. da Silva, "A systematic review of game design methods and tools," In *International Conference on Entertainment Computing,* pp. 17-29, Springer, Berlin, Heidelberg, 2013.

[42] P. Heinrich and G. Schwabe, " Communicating nascent design theories on innovative information systems through multi-grounded design principles," In *International Conference on Design Science Research in Information Systems*, pp. 148-163, Springer, Cham, 2014.

[43] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *Journal of management information systems*, *24*(3), 45-77, 2007.

[44] M. Rosewater, "Timmy, Johnny, and Spike Revisited. *Making Magic*, *18,"* [ONLINE] Available from: https://magic.wizards.com/en/articles/archive/making-magic/timmy-johnny-and-spike-revisited-2006-03-20, 2006, accessed May 2020.

[45] R. M. Ryan and E.L. Deci, "Intrinsic and extrinsic motivations: Classic definitions and new directions," *Contemporary educational psychology*, *25*(1), 54-67, 2000.

[46] J. Juul and M. Norton, "Easy to use and incredibly difficult: on the mythical border between interface and gameplay," In *Proceedings of the 4th international conference on foundations of digital Games*, pp. 107-112, 2009.

[47] R. Bartle, "Hearts, clubs, diamonds, spades: Players who suit MUDs," *Journal of MUD research*, *1*(1), 19, 1996.

[48] R. Hunicke, M. LeBlanc, and R. Zubek, "MDA: A formal approach to game design and game research," In *Proceedings of the AAAI Workshop on Challenges in Game AI*, Vol. 4, No. 1, p. 1722, 2004.

[49] N. Lazzaro, "Why we play games: Four keys to more emotion in player experiences," In *Proceedings of GDC*, 2004.

[50] M. Csikszentmihalyi, *Flow: The psychology of happiness*. Random House, 2013.

[51] J. Hamari and J. Tuunanen, "Player types: A meta-synthesis", 2014.

[52] B. Stewart, "Personality and play styles: A unified model," *Gamasutra, September*, *1,* 2011, [ONLINE] Available from: *http://www. gamasutra. com/view/feature/6474. php.*, accessed May 2020.

[53] C. Liu, P. Agrawal, N. Sarkar, and S. Chen, "Dynamic difficulty adjustment in computer games through real-time anxiety-based affective feedback," *International Journal of Human-Computer Interaction*, *25*(6), 506-529, 2009.

[54] R. Khaled, "Muse-based game design," In *Proceedings of the Designing Interactive Systems Conference*, pp. 721-730, 2012.

[55] J. VandenBerghe, "Engines Of Play Case Study: Relic Entertainment," 2017, [ONLINE]. Available from: https://www.darklorde.com/engines-play-case-study-relic-entertainment/, accessed May 2020.

[56] C. Hodent, *The gamer's brain: How neuroscience and UX can impact video game design*, CRC Press, 2017.

[57] C. Treadaway, A. Taylor, and J. Fennell, "Compassionate design for dementia care. *International Journal of Design Creativity and Innovation*, *7*(3), 144-157, 2019.

[58] D. A. Norman, "Cognitive engineering. *User centered system design*, *31*, 61, 1986.

[59] M. Rajanen and J. Nissinen, "A Survey of Game Usability Practices in Northern European Game Companies," IRIS: Selected Papers of the Information Systems Research Seminar in Scandinavia. Issue Nr 6, 2015, Paper 8. http://aisel.aisnet.org/iris2015/8

[60] C. H. Vuong, T. Ingalls, and J. J. Abbas, "Transforming clinical rehabilitation into interactive multimedia," In *Proceedings of the 19th ACM international conference on Multimedia*, pp. 937-940, 2011.

[61] B. Dewar, E. Adamson, S. Smith, J. Surfleet, and L. King, "Clarifying misconceptions about compassionate care," *Journal of Advanced Nursing*, *70*(8), 1738-1747, 2014.

[62] R. Bernhaupt, *User Experience Evaluation Methods in the Games Development Life Cycle*. Springer International Publishing, Cham, 1–8, 2015.

[63] T. De la Hera, E. Loos, M. Simons, and J. Blom, "Benefits and factors influencing the design of intergenerational digital games: a systematic literature review," *Societies*, *7*(3), 18, 2017.

[64] B. Suits, *The Grasshopper-: Games, Life and Utopia*. Broadview Press, 2014.

[65] B. Sutton-Smith, *The ambiguity of play*. Harvard University Press, 2009.

[66] C. Crawford, *Chris Crawford on game design*, New Riders, 2003.

[67] F. Mäyrä, *An introduction to game studies*, Sage, 2008.

[68] Wizard of the Coast, *Magic: The Gathering*, [Card Game] US.Wizard of The Coast, 1993.

[69] Niantic, *Pokemon Go*, [Mobile Game], Ninantic, 2016.

Techniques and Methodologies for Measuring and Increasing the Quality of Services: a Case Study Based on Data Centers

Martin Zinner*, Kim Feldhoff*, Michael Kluge*, Matthias Jurenz*, Ulf Markwardt*, Daniel Sprenger*,
Holger Mickler*, Rui Song†, Björn Gehlsen*, Wolfgang E. Nagel*
* Center for Information Services and High Performance Computing (ZIH)
Technische Universität Dresden
Dresden, Germany
E-mail: {martin.zinner1, kim.feldhoff, michael.kluge, matthias.jurenz}@tu-dresden.de,
{ulf.markwardt, daniel.sprenger, holger.mickler}@tu-dresden.de,
{bjoern.gehlsen, wolfgang.nagel}@tu-dresden.de
† Technical Information Systems
Technische Universität Dresden
Dresden, Germany
E-mail: rui.song@tu-dresden.de

***Abstract*—As data centers become increasingly complex and deliver services of high importance, it is very important that the quality of the delivered services can be objectively evaluated and can fulfill the expectations of the customers. In this paper, we present a novel, general, and formal methodology to determine and improve the Quality of Services (QoS) delivered by a data center. We use a formal mathematical model and methodology in order to calculate the overall indicator of the service quality and discuss methods of improving the QoS. Since the considerations were conceived and results have been proved in a formal model, the considerations and results also hold in a more general case. We discuss the pros and cons of the Continuous Change Strategy and analyze the Customer Dissatisfaction (CD) concepts. We show that CD is not the opposite of Customer Satisfaction (CS), but it can be used in a meaningful way to estimate CS. We introduce the queueing model and use the operation curve and the flow factor to improve the performance of data centers.**

Keywords*–Quality of Services; QoS; Performance data center; Little's Law; Kingman's equation; Flow factor; Operating curve management; Customer satisfaction; Key performance indicators; Customer satisfaction; Customer dissatisfaction; Continuous change strategy; Continuous delivery.

I. INTRODUCTION

A. Motivation and Short Overview

Nowadays, services of high quality of data centers are indispensable for the good functioning of a company or a research institute. However, due to the advanced digitalization, data centers are becoming more and more complex and difficult to manage [1]. According to a survey of Symantec [2], the main reasons are the raise of the Cloud Computing and the Virtualization. Basically, such complex infrastructures are more error-prone and require more maintenance efforts than simple ones. Thus, it is of crucial importance to measure the Quality of Services (QoS) provided by a data center, in order to detect which components / services are low performers and should be improved. This way, measuring the QoS also avoids service degradations. Services which underperform can be detected and measures can be taken (like relocation of resources) such that these services will perform better again. An optimized usage of the available resources does not only improve the QoS and thus, the image of the service provider, but also helps to save costs.

Furthermore, Butnaru [3] states that "quality has become a strategic element in companies dealing with services because it determines competitiveness at its highest level". Thus, by measuring and improving the QoS provided by companies / research institutes, the service providers can improve their ranking when compared to the competition.

Estimating the QoS of a data center is a complex endeavor. On the one hand, there are objectively measurable indicators like the duration and number of unplanned down times. On the other hand, the customer satisfaction has very important subjective components, which should not be neglected. Thus, if a customer has full confidence in the technical skills, seriousness, and professionalism of the operating staff, then his/her attitude is permissive and indulgent regarding possible malfunctions. For example, let us consider the scenario that a service has an unplanned downtime. If the operating staff can predict the time when the resumption of the service will occur with satisfying accuracy, the impression of the customer regarding the service provider will be very good. Otherwise, the customer will assume that the service provider does not have his/her processes under control and a failure of the system will sooner or later occur.

In this paper, we will focus on the perspective from the data center side. We will define and make use of different metrics in order to be able to establish objective criteria which characterize the Quality of Services and the performance of a data center.

B. Main Challenges and Objectives

If a customer is asked about the quality of the services of a data center he or she usually will answer: Yes, quality is good, but it could be better. This answer only describes the subjective perception of the customer. Our aim is to go further. Thus, searching for a positive response to the questions "Is the QoS measurable and if this is the case, how?" is one of the main challenges, we had to accept and take up.

Establishing and choosing meaningful performance indicators form the basis for improving the QoS.

The challenges described above lead us to the following main objectives:

1) Receive responses to the question whether the QoS is quantifiable / metrisable or not, i.e., whether the QoS can be expressed numerically in a reasonable non trivial way, such that this number is independent of the subjective perception of humans.
2) Establish a general approach on the modalities to quantify the QoS such that a single indicator (or only very few) expresses the service quality of the service provider.
3) Find possibilities to improve the QoS and implicitly the performance of the service provider.

C. Outline

The remainder of the paper is structured as follows: Section II gives a short overview of the state of the art and detail some difference of our approach. Section III introduces the proposed strategy for measuring the QoS, first in an informal way, afterwards formalized by introducing a mathematical model. Different metrics are defined and used in order to be able to establish objective criteria which characterize the quality of the services and the performance of a data center. Section IV introduces the formal queueing model, makes the connection to the models used in practice, and discusses modalities to improve the performance of a data center by using the operation curve. In order to be able to balance the performance of the services of different departments of a data center, a formula to calculate the flow factor of the data center out of the flow factor of each department, is established. Section V covers additional strategies to improve the QoS of a data center as the Continuous Change Strategy (CCS) and the Customer Dissatisfaction (CD) concept. Section VI gives some details of a use case and finally, Section VII concludes this paper and sketches the future work.

II. Related work

An important part of the existing approaches for quality improvement focus merely on the QoS from the user perspective – established through questionnaires (e.g., SERVQUAL and/or SERVPERF) [4]) – and on the discrepancies between the user perception and the user expectation of the QoS.

Most approaches concerning the measurement of QoS have tended to avoid the use of pre-defined objective performance indicators and focus instead on the relationship between what consumers expect from a particular service and what they actually get [5]. The conclusion [4] is that customer satisfaction with services or perception of QoS can be viewed as confirmation or disconfirmation of customer expectations of a service offer. The role of emotions in customer-perceived service quality is analyzed [6] by widening the scope of service quality, i.e., by focusing on dimensions beyond cognitive assessment.

We concentrate our study primarily on the service provider perspective by using metrics to characterize the QoS and subsequently establish strategies on how those metrics can be combined together to generate a unique indicator, which characterizes the overall performance of the service provider.

Measuring and ranking service quality has been an issue for study for decades [5], whereby the difficulties lied in the development of the most suitable method of measurement. Approaches to the measurement of QoS are based on the analysis of the relationship between customer expectation of a service and their perceptions of its quality. Indices to provide measures of expectation, perceptions and overall satisfaction from the customer side are set up [5] and compared.

In [7], the authors report the insights obtained in an extensive exploratory investigation of quality in four business (retail banking, credit card, security brokerage, and product repair and maintenance) by developing a model of service quality. The most important insight obtained from analyzing the executive responses is the following: “A set of key discrepancies or gaps exists regarding executive perception of service quality and the tasks associated with service delivery to consumers. These gaps can be major hurdles in attempting to deliver a service which consumers would perceive as being of high quality”.

Metrics in order to establish the QoS have been used for example, by the Systemwalker [8], which supports “Information Technology Infrastructure Library” (ITIL) based IT service management. The focus in [8] is on the service delivery area, such as capacity, availability, and service level management. The composition of metrics is outside the scope of the Systemwalker.

In [9], a framework for the evaluation of QoS for Web Services within the OPTIMACS project is presented, such that Service Level Agreements (SLAs) are established in order to calculate / guarantee the QoS, then the properties are normalized by using statistical functions. The goal is to obtain a final Quality grade, which allows to rank the services. Finally, aggregation is performed using weighted sum of the different quality items.

As a final note, the studies regarding the normalization and composition of metrics considered for QoS for Web Services are straightforward and are based on statistics (minimum, maximum, mean, standard deviation, Z-score) [9], the committed SLA time provides the QoS level. The metrics used to measure the QoS of a data center are so diverse that a case-by-case approach is necessary to determine the normalization and composition strategy. Moreover, statistical values as above are generally not a priori known for unconverted metrics such as “cycle time”, etc.

III. Measuring the QoS

We describe the general strategy how to measure the QoS in an informal way in Subsection III-A and formalize this strategy in Subsection III-B.

A. Description of the Strategy

According to ITIL [10] (and similar), the (incomplete) list of processes comprises the following managements:

- “incident management”,
- “problem management”,
- “information security management”,
- “service level management”,

- “change management”,
- “project management”, and
- “release and deployment management”.

A list of metrics is specified for each process according to ITIL [10] and / or to the “Key Performance Indicator Library" (KPI Library) [11], see [12] regarding developing, implementing and using KPIs. Some of the most important metrics for the ITIL process “incident management” are given in the following:

- “Total number of incidents”,
- “Number of repeated incidents, with known resolution methods”,
- “Number of incidents escalated which were not resolved in the intended resolution time”,
- “Average cycle time associated to the subsequent responses”, i.e., including the average cycle time to resolve the incident,
- “Average waiting time from user side associated to the subsequent responses”,
- “Average work effort for resolving the incident associated to the subsequent responses”,
- “Average time between the occurrence of an incident and its resolution”,
- “Total number of incidents resolved within service level agreement (SLA) time divided by the total number of incidents”.

In order to establish objective criteria for measuring the QoS, it is not sufficient to consider simply one metric. Indeed, different metrics have to be combined. The following example illustrates this issue.

The metric “Total number of incidents” is a revealing metric regarding the performance of a service provider. Of course, this metric is important for reporting per se, as a non anticipated sharply increasing trend can be the cause for major concerns. Another important metric is "the average cycle time to solve an incident". If the metric “Total number of incidents” is increasing, but in the mean time the “Average cycle time to solve an incident” is decreasing, the balance is restored and the service provider will not face a total collapse of the service.

Hence, composition rules for metrics are needed, such that indicators that characterize the health of the services, can be established. Since we cannot directly compare the different metrics, we transform / normalize the metrics using relative values. By dividing the “Total number of incidents” by an artificially generated “Maximum number of incidents supported”, we receive a relative value between 0 and 1. Unfortunately, the value 1 is the worst value you can ever get. In order to circumvent this impediment, we subtract 1 and change the sign. Using the same considerations (by defining the “Minimum average cycle time”) analogue relative value for the cycle time can be established. In this case, this new indicator is directed in the sense that the best cycle time is achieved when this value is equal to 1. This example is just to illustrate the technique. One may argue that an increase of the indicator value “Average waiting time of the incidents during processing” also indicates a congestion.

Thus, in order to combine different metrics, we will normalize them to the range $[0; 1]$ in such a way that the lowest value correspond to the poorest quality, the highest value to the best quality. Once, all the relevant metrics of a process are normalized, we can proceed with the composition such that for each process a single, composed metric is established.

The composed metric should also take values between 0 and 1, such that a greater value implies a better QoS. An example of a straightforward composing strategy is to establish weights for each metric, such that the sum of all weights is equal to 1 and important metrics have bigger weights. Hence, the decisive metrics are much better considered. Of course for practical purposes, we can define groups of incidents having the same importance and accordingly appropriate distribution functions (linear, exponential, etc.). The calculation of the associated weights is then immediate.

Normally, explicitly defining importance grouping and distribution functions is not always necessary. We can set up priority strategies regarding the QoS. As an example, under some circumstances, a fast but not necessarily very detailed answer is more helpful for IT professionals, who can elaborate the details themselves. In other cases, detailed and very accurate answers are necessary, especially for customers with little or no experience. Then, customers could return the ticket of the incident (e.g., if the answer is not accurate enough) and ask for more information and assistance.

Hence, the development of an appropriate strategy for the quality improvement is essential, in some cases this strategy can be even customer dependent. For example, we can improve the quality:

- by improving only the accuracy of the responses, or
- by reducing only the response times, or
- by minimizing a metric which takes both accuracy and the response time into account.

In accordance with the improvement strategy, the grouping of metrics regarding their performance is more or less straightforward and easy to follow.

In effect, we can establish for each process a unique (abstract) indicator, which characterizes the quality of the process such that a greater value means better quality of the process according to the improvement strategy as above. The absolute value of this indicator has no particular interpretation, only the increment or decrement of this value in time is significant.

Same considerations using the indicators established for the processes lead to a unique indicator of the QoS for the whole service provider, i.e., the data center. By evaluating the time behavior of this indicator and / or the component indicators we can have a good overview which process and / or metrics performed better or worse.

This unique indicator can be deployed for example, on daily bases, such that the performance of the service provider can be easily followed and appropriate measures can be taken if performance degradation occurs. Moreover, even if the overall unique indicator has improved in value, there can be some components, whose performance has degraded. By setting up appropriate Graphical User Interfaces (GUIs), and appropriate colors (for example, red for degradation and green for improvement) the deviation with respect to the previous day can be visualized.

The only process through which the customers interact with the data center as the service provider is through the "incident management", the performance of the other processes is practically hidden for the regular customer. In order to improve the "incident management" we will analyze the impact of some important processes on the "incident management".

An important direct impact on the "incident management" has the "problem management" in the sense that by a very efficient "problem management" the number of repetitive incidents or the time to solve the repetitive incidents can be dramatically reduced. For this, each incident should be correctly assigned to the appropriate issue, have a correct and exhaustive root analysis, such that the causes of the incident are unambiguously elucidated. It seems a bit of common sense that all the detailed information regarding the incident including good ways of searching, finding, and retrieving the information should be stored in an appropriate knowledge database. Next, the probability of recurring should be estimated and if necessary, appropriate measures should be taken in order to avoid the next occurrence of the same incident.

Proactive methods are very efficient to avoid the occurrence of incidents, e.g., improving "change management", "release and deployment management", etc. By significantly reducing the impact of new releases on the services, the peaks on the QoS can be significantly reduced.

The long term personal experience of the first author – working as a software engineer at Infineon / Qimonda – is that the behavior of the information technology systems – including also the developer and maintenance staff – was almost optimal under steady state conditions. We did not have any theoretical explanation for this behavior, we only have noticed that any larger deviation (i.e., non stable state) from the steady state required unpredictable efforts to return to a stable environment.

Based on the above, we present the concept of Continuous Change Strategy (CCS) and discuss the advantages and disadvantages later on. The basic idea behind CCS is that a major release is split into a larger number of small releases, which are put into production as soon as they have passed the appropriate acceptance tests. This way, the major release change is accomplished as soon as the last minor release has been deployed into production.

B. *Formalization of the Strategy*

We will formalize the strategy [13], [14] proposed in Subsection III-A by introducing a mathematical model in order to use the advantages of the rigor of a formal approach over the inaccuracy and the incompleteness of natural languages.

Let $\mathcal{A}$ be an arbitrary set. We notate by $2^{\mathcal{A}}$ the power set of $\mathcal{A}$, i.e., the set of all subsets of $\mathcal{A}$, including the empty set and $\mathcal{A}$ itself, and the cardinality of $\mathcal{A}$ by $card(\mathcal{A})$.

We use a calligraphic font to denote index sets. We denote by $\mathcal{S} := \{S_i \mid i \in \mathscr{S}$ and S_i is a service$\}$ the finite set of the services. Analogously, we denote by $\mathcal{P} := \{P_i \mid i \in \mathscr{P}$ and P_i is a process$\}$ the finite set of processes and by $\mathcal{T} := \{[t_1, t_2] \mid t_1$ and t_2 are points in time, such that $t_1 \leq t_2\}$ time intervals.

A metric M is a measurement that monitors progress towards achieving the targeted objectives. We denote by $\mathcal{M} := \{M_i \mid i \in \mathscr{M}$ and M_i is a metric$\}$ the finite set of metrics. Generally speaking, a metric M is defined for an environment containing subsets of $\mathcal{S}$ and $\mathcal{P}$.

For example, let us define the ratio between the "total number of incidents with known resolution method" and the "total number of incidents". Depending on the strategic orientation of the company, different goals can be pursued. On the one hand, having for most of the incidents corrective measures in place can be a targeted objective, one the other hand, avoiding repetitive incidents is crucial for the economic success of companies like fabs running 24x7 continuous manufacturing operations. A mixed strategy (for example, 10% known errors) can be also targeted. Hence, the scope of a metric is most of the time business oriented.

In order to be able to compare and compose different metrics in a reasonable way, we introduce the value of a metric such that it is greater or equal 0 and lower or equal 1. A greater value of the metric means a closer value to the targeted business objectives. Formally, the range of values of the *possible* business values, including the targeted ones is $2^{\mathbb{R}}$. Hence, the progress towards achieving the targeted Business Objectives (BO) can be represented as a function.

$$\begin{aligned} BO : \mathcal{M} \times \mathcal{P} \times \mathcal{S} &\to BusinessObjectives, \\ (M, P, S) &\mapsto BO(M, P, S). \end{aligned}$$

Analogously, the value (V) of a metric is represented as:

$$\begin{aligned} V : \mathcal{M} \times \mathcal{P} \times \mathcal{S} \times \mathcal{T} &\to [0, 1], \\ (M, P, S, [t_1, t_2]) &\mapsto V(M, P, S, [t_1, t_2]). \end{aligned}$$

A greater value for $V(M, P, S, [t_1, t_2])$ means a closer value to the targeted business objectives for (M, P, S). The definition above highlights the fact that the same metric can have different business objectives and definition (values) depending on the environment (services and/or processes) it is used.

We illustrate the above considerations based on a simple example and consider the "average cycle time" of the incidents. The business demands short cycle times for all departments. In order to be able to compare the cycle times of different departments, we determine the minimal cycle time (i.e., the theoretical cycle time needed if there are no unplanned down times, etc.) and assign the ratio of minimal cycle time to the cycle time as the value of the metric. Hence, the performance of the different departments regarding the same metric (i.e., cycle time) can be easily compared, on the assumption that the respective minimal cycle time has been evaluated correctly.

Our aim is to establish a single indicator for the service performance (i.e., the QoS) of the service provider. In order to evaluate the performance of the different metrics of the same process (for example, ITIL process), we set up a methodology to compose the different metrics in a reasonable way, such that the new metric (indicator) outlines the performance of the investigated process.

In order to simplify the notation, we will notate in the following the value of a metric M by $V(M)$, meaning that the metrics involved are defined on the same environment and the same time interval.

Definition III.1 (Composition of metrics) *Let $\mathfrak{M} := \{M_i | \, i \in \{i_1, i_2, \ldots, i_k\} \subseteq \mathcal{M}\}$ a subset of $\mathcal{M}$. We define*

$$\begin{aligned} COMP : 2^{\mathcal{M}} &\to \mathcal{M}, \\ \mathfrak{M} &\mapsto COMP(\mathfrak{M}), \end{aligned}$$

such that there is an aggregation function AGG

$$\begin{aligned} AGG : 2^{\mathcal{M}} &\to [0,1], \\ V(COMP(\mathfrak{M})) &\mapsto AGG(V(M_{i_1}), V(M_{i_2}), \ldots, V(M_{i_k})) \end{aligned}$$

and

$$\begin{aligned} &v^1_{i_1} \leq v^2_{i_1}, v^1_{i_1} \leq v^2_{i_1}, \ldots, v^1_{i_k} \leq v^2_{i_k} \\ &\quad \Rightarrow AGG(v^1_{i_1}, v^1_{i_2}, \ldots, v^1_{i_k}) \leq AGG(v^2_{i_1}, v^2_{i_2}, \ldots, v^2_{i_k}) \end{aligned}$$

Except for the case of trivial aggregations, the composition generates a new metric out of known ones.

In order to keep our notation simple and straightforward, we will not make any distinction in the formal representation of the initial metrics and those obtained by consolidation. Hence, $\mathcal{M}$ contains the initial metrics as well as the consolidated ones. Therefore, a consolidated metric can be finally set up for the entire service. We note:

Lemma III.2 (Composition properties) *Let $\mathfrak{M} := \{M_i | \, i \in \{i_1, i_2, \ldots, i_k\} \subseteq \mathcal{M}\}$ a subset of $\mathcal{M}$ arbitrarily chosen. Then, $COMP(\mathfrak{M})$ is a metric, i.e., fulfills the following properties:*

a) $0 \leq V(COMP(\mathfrak{M})) \leq 1$,

b) *A greater value for $V(COMP(\mathfrak{M}))$ means a closer value to the targeted business objectives for this metric.*

Hint *These properties are a direct consequence of Definition III.1.* ■

Next, we give a small example to illustrate the aggregation strategies. Let $\mathfrak{M} := \{M_{i_1}, M_{i_2}, \ldots, M_{i_k}\}$ be a subset of $\mathcal{M}$. We suppose that the value of the new characteristic $COMP(\mathfrak{M})$ is a linear combination of the values of the components, i.e.,

$$V(COMP(\mathfrak{M})) := \sum_{i=i_1}^{i_k} \alpha_i \cdot V(M_i)$$

with weights $\alpha_i > 0 \; \forall i \in \{i_1, i_2, \ldots, i_k\}$ and $\sum_{i=i_1}^{i_k} \alpha_i = 1$. If the value of α_i is high, then the metric M_i within $COMP(\mathfrak{M})$ is important. In practice, it suffices to build weight groups $\{G_i | \, i \in \{i_1, i_2, \ldots, i_l\}\}$ out of $\mathfrak{M}$ such that each $M \in \mathfrak{M}$ belongs to a group G_i and all $M_j \in G_i$ are equally weighted. Furthermore, a weighting function W can be set up, such that all α_i can be explicitly determined, for example, set

$$k_i := \frac{\alpha_i}{\alpha_{i+1}} \text{ for } i \in \{i_1, i_2, \ldots, i_{l-1}\}.$$

The values k_i can be regarded as the "ratio of relevancy" of the corresponding metrics.

IV. Improving the QoS

In the last section, we proposed a strategy how to measure the QoS of a data center. In this section, we will establish metrics controlling the performance of a data center. Thus, we can determine in which cases the service of a data center collapses or the QoS substantially degrades.

A. Queueing Model and Basic Metrics

We model the processing line of a data center by introducing a queueing model and give some basic definitions related to it. In order to keep the presentation accessible and avoid technical complications, we will maintain our model as simple as possible. It is the task of the practitioners to map the real world onto this model according to their needs. We will analyze the entire processing line as well as subsystems of it.

A *queueing system* consists of discrete objects, termed *units* or *items* that arrive at some rate to the system. Within the system, the units may form one or more queues and eventually be processed. After being processed, the units leave the queue.

The finest granularity in our model is *unit*, *step*, *time stamp*, *section* and *classification*. For example, in practice, the unit can be a ticket, the section can be an employee of the service center, a group of employees having the same profile or a specific section of the service center, etc. The classification is the finest attribute which characterizes the unit (like bug, disturbance, project, etc.) and it can be distinguished in the processing phase.

In our model the unit enters the system (service center), is processed according to the specifications and leaves the system. The step is the finest abstraction level of processing which is tracked by the reporting system. When the material unit u enters the system, it is assigned to a classification c. This assignment remains valid till the unit u leaves the system. We will analyze the entire processing line as well as subsystems of it.

We denote by S the set of all steps of the processing line, by U the set of the units that entered the system, and by T the (ordered and discrete) points in time when events may occur in the system. Since we are merely interested in daily calculations, we will set D as the set of all points in time belonging to a specific day D, i.e., $D := \{t \in T | \text{ t belongs to day D}\}$.

Let $s \in S$ and $u \in U$. We denote by $TrInT_s(u)$ the *track in time* of u, i.e., the point in time when the processing of unit u is started at step s. Analogously, $TrOutT_s(u)$ is the *track out time* of u, i.e., the point in time when the processing of unit u has been finished at the step s.

We assume that for a step s, the function $succ_s(u)$, which identifies the succeeding step of s for the unit u is well defined. Analogously, we assume that the history of the production process is tracked, so the predecessor function $pred_s(u)$ of each step s is well defined. For formal reasons we set $succ_s(u) := s$ for the last step on the route and $pred_s(u) := s$ for the first step on the route.

By *cycle time* (CT), we generally denote the time interval a unit or a group of material units spent in the system / subsystem [15]. We do not make any restrictions on the *time unit* we use, but are merely interested on daily calculations.

For formal reasons, – in order to be able to calculate average values – we denote by $24h$ the cardinality of an arbitrary day D. For $t \in T$ we denote by $t \pm 24h$ the point in time t shifted forward or backwards by 24 hours.

We assume that events in the system are repeated on a daily basis, i.e.,

$$\begin{aligned} &\forall u \in U \text{ and } \forall s \in S : \mathit{TrInT}_s(u) = t \\ &\implies \exists v \in U : \mathit{TrInT}_s(v) = t + 24h \text{ and} \\ &\qquad \mathit{TrOutT}_s(v) = \mathit{TrOutT}_s(u) + 24h \end{aligned}$$

and

$$\begin{aligned} &\forall u \in U \text{ and } \forall s \in S : \mathit{TrInT}_s(u) = t \\ &\implies \exists w \in U : \mathit{TrInT}_s(w) = t - 24h \text{ and} \\ &\qquad \mathit{TrOutT}_s(w) = \mathit{TrOutT}_s(u) - 24h. \end{aligned}$$

Under a *stable system* we mean a system according to the conditions above.

In practice, systems pass through a ramp up phase such that the above conditions are eventually reached, i.e., $\exists t_b \in T$ such that the above conditions are satisfied for all $t > t_b$. For our investigations, it is sufficient that the systems reach the stable state after some time (*eventually stable systems*). For reasons of a simple notation, we will use the term *stable system* or *system in a stable state*. However, the statements of this work are also valid for eventually stable systems.

The *raw process time / service time* of unit $u \in U$ related to step $s\ inS$ is the minimum processing time to complete the step s without considering waiting times or down times. We denote the raw process time of unit u related to step s by $RPT_s(u)$.

Let $u \in U$, let $\{s_1, s_2, \ldots, s_n\} \subset S$ be the complete list of steps according to the processing history to process unit u. Let $RPT_{s_i}(u)$ be the raw process times of unit u related to step s_i for all $i = 1, 2, \ldots, n$. Then the raw process time of unit u can be represented as follows:

$$RPT(u) = \sum_{i=1}^{n} RPT_{s_i}(u).$$

The work in progress is defined as the inventory at time $t \in T$ and will be denoted by $WIP(t)$. If the work in process is used in connection with Little's Theorem then it denotes the average inventory for a given period of time. We use the notation $avgWIP$ instead of WIP to denote the average inventory.

We denote by Th the throughput of the material units by. Usually, we consider the daily throughput and refer to it as Th^D for a specific day D.

The cycle time of a unit $u \in U$ spent at a step $s \in S$ in the system can be represented as:

$$CT_s(u) := \mathit{TrOutT}_s(u) - \mathit{TrOutT}_{pred_s(u)}(u).$$

Let $\{u_1, u_2, \ldots, u_n\}$ be the set of units that were processed at step s on a specific day $D \subset T$ i.e., $\forall i \in \{1, 2, \ldots, n\}\ \exists t_i \in D$ such that $\mathit{TrOutT}_s(u_i) = t_i$. Then, the average cycle time $avgCT_s^D$ the units u_i spent in the system at step s on a specific day D can be represented as:

$$avgCT_s^D = \frac{1}{n} \cdot \sum_{i=1}^{n} CT_s(u_i).$$

For $t \in T$, $u \in U$ we define the indicator function 1_s at a process step $s \in S$ as follows:

$$\begin{aligned} &1_s : U \times T \to \{0, 1\}, \\ &(u, t) \mapsto 1_s(u, t) := \begin{cases} 1 & \text{if } t \geq \mathit{TrOutT}_{pred(s)}(u) \text{ and} \\ & \quad t < \mathit{TrOutT}_s(u), \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Throughout this work we assume that T is discrete, i.e., units arrive and depart only at specific points in time, since the time is usually measured in seconds or milliseconds.

Lemma IV.1 (Representation of average inventory) *The average inventory $avgWIP_s^D$ for a process step $s \in S$ on a specific day D can be represented as follows:*

$$\begin{aligned} avgWIP_s^D &= \frac{1}{\text{card}(D)} \cdot \sum_{t \in D} \sum_{u \in U} 1_s(u, t) &(1) \\ &= avgCT_s^D \cdot Th_s^D. &(2) \end{aligned}$$

By interchanging the order of summation, we receive an expression for WIP_s^D, which is much easier to calculate in practice.

Hint *Let $U_{n,D}$:= $\{u_1, u_2, \ldots, u_n\}$ be the set of units that left the step s on a specific day $D \subset T$, i.e., $\forall i \in \{1, 2, \ldots, n\}\ \exists t_i \in D$ such that $\mathit{TrOutT}_s(u_i) = t_i$. Then, in stable systems the following relation holds:*

$$\begin{aligned} avgWIP_s^D &= \frac{1}{\text{card}(D)} \cdot \sum_{t \in D} \sum_{u \in U} 1_s(u, t) \\ &= \frac{1}{\text{card}(D)} \cdot \sum_{t \in T} \sum_{u \in U_{n,D}} 1_s(u, t). \end{aligned}$$

By interchanging the order of summation and considering that for i $\in \{1, 2, \ldots, n\}$ the average cycle time (measured in days) of the material unit u_i at step s is given by:

$$\begin{aligned} &avgCT_s(u_i) \\ &:= \frac{1}{\text{card}(D)} \cdot \left(\mathit{TrOutT}_s(u_i) - \mathit{TrOutT}_{pred(s)}(u_i)\right) \\ &= \frac{1}{\text{card}(D)} \cdot \sum_{t \in T} 1_s(u_i, t). \end{aligned}$$

Thus, we get:

$$avgWIP_s^D = avgCT_s^D \cdot Th_s^D.$$

Since in stable systems the variables above do not depend on the day chosen for their calculation, Little's Theorem follows. The consideration above do not hold in steady state systems used by Stidham and Sigman (see [16] or [17]). ■

Remark IV.3 (Case: Set of points in time is continuous) *Actually, in theoretical models it is not necessary to consider a discrete set T in order to be able to calculate $avgWIP_s^D$. Let Σ_U be the discrete σ-algebra on U (i.e., the power* *set 2^U of U). Let μ be the counting measure on Σ_U, i.e., $\mu(\mathcal{U}) := |\mathcal{U}|$ for $\mathcal{U} \in \Sigma_U$. Then, (U, Σ_U, μ) is a measure space. For $T \subset \mathbb{R}_+$ let Σ_T be the σ-algebra of all Lebesgue measurable sets on T and let λ the usual Lebesgue measure on T. Analogously (T, Σ_T, λ) is also a measurable space. Since both spaces are σ-finite, the product measure $\mu \otimes \lambda$ is well defined and for $\mathcal{U} \subset U$ and $\mathcal{T} \subset T$ the equality $\mu \otimes \lambda(\mathcal{U} \times \mathcal{T}) = \mu(\mathcal{U}) \cdot \lambda(\mathcal{T})$ holds. Since 1_s is a simple function (i.e., a finite linear combination of indicator functions of measurable sets) it is $\Sigma_U \times \Sigma_T$ measurable. Then, as expected $\mathrm{card}(D) = \int_D \mathrm{d}\lambda(t) = 24h$ and the theorem of Fubini-Tonelli gives:*

$$\begin{aligned} avgWIP_s^D &= \frac{1}{\mathrm{card}(D)} \cdot \int_D \int_U 1_s(u,t)\mathrm{d}\mu(u)\mathrm{d}\lambda(t) \\ &= \frac{1}{\mathrm{card}(D)} \cdot \int_{U \times D} 1_s(u,t)\mathrm{d}(\mu \otimes \lambda)(u,t) \\ &= \frac{1}{\mathrm{card}(D)} \cdot \int_U \int_D 1_s(u,t)\mathrm{d}\lambda(t)\mathrm{d}\mu(u). \end{aligned}$$

The last integral is much easier to evaluate.

In stable systems the value $avgWIP_s^D$ does not depend on the specific day D that was considered for the calculation.

B. Expected Inventory

Next, we define one of the relevant metric for bottleneck control and present formulas to calculate them.

$WIP24_{s,c}(t)$ denotes the inventory which is expected in the next 24 hours at a specific step $s \in S$, classification c and $t \in T$. Usually, $WIP24_{s,c}(t)$ at midnight is considered. In this case, we will omit the time constraint and use the notation $WIP24_{s,c}$.

Let us suppose $\{s_1, s_2, \dots, s_n\}$ is the planned (ordered) list of steps as provided by the route for the classification c. There are of course different strategies to estimate $WIP24_{s_l,c}(t)$ for a specific $l \in \{1, 2, \dots, n\}$. One alternative supposes that the units moves across the line as planned by the route. Let $refCT_{s_i,c}$ be the target cycle time to process the unit at the step $s_i \in S$, let $WIP_{s_i,c}(t)$ be the inventory at the step s_i for the classification c and time $t \in T$. For l determine $j := min(k : k \leq l)$ such that $\sum_{k \leq i \leq l} refCT_{s_i,c} \leq 24h$. Then the expected inventory can be written as follows:

$$WIP24_{s_l,c}(t) = \sum_{j \leq i \leq l} WIP_{s_i,c}(t).$$

Most of the time, the unit is not processed according to the specifications (route), reworks or alternative processing strategies are necessary. In this case, the formula as above does not hold, and other more complex approaches are necessary.

C. Little's Theorem

In the following, we will introduce Little's Theorem [18] [19]. Little's Theorem which is mostly called *Little's Law* is a mathematical theorem giving a rather simple relation between the average cycle time, the throughput, and the average work in process in the system. It will be used later on for calculating the flow factor and thus, controlling the performance of the data center. The relation of Little's Theorem is valid if some convergence criteria are fulfilled and if the underlying system is in steady state and non-preemptive. The latter means that the properties of the system are unchanging in time, there are no interrupts and later resumes. In many systems, steady state is not achieved until some time has elapsed after the system is started or initiated. In stochastic systems, the probabilities that some events occur in the system are constant. The result is entirely independent of the probability distribution involved and hence it requires no assumption whether the units are processed in the order they arrive or the time distribution they enter or leave the system.

We give now a formal definition for Little's formula. Our explanation is based on [17] slightly modified to use our notations. We consider the queueing system above where – unlimited but countable – units arrive, spend some time in the system, and then leave. Material units enter at most once the system, i.e., units that left do not enter the system again. Let $T := \{t_i | \, i \in \mathbb{N}\}$ be the countable set of points in time when those events occur. At any point in time $t \in T$ at most a finite number of units enter or leave the system. Let u_n denote the unit which enters the system at the time t_n^e. Upon entering the system, u_n spends CT_n time units in the system (the cycle time of u_n) and then leaves the system at time $t_n^d = t_n + CT_n$. The departure times are not necessary ordered in the same way as the enter times. This means that we do not require that the units leave the system in the same order as they arrived. Let $1_{u_i}^e(t) := 1$ if $t_i \leq t$ and 0 else. We denote by $N^e(t)$ the number of units which entered the system until time t, i.e.,

$$N^e(t) = \sum_{i=1}^{\infty} 1_{u_i}^e(t).$$

Analogously, we denote by $N^l(t)$ the number of units which have left until time t. Let $L(t)$ be the total number of the units in the system by time t. A unit u_n is in the system at time t if and only if $t_n \leq t < t_n + CT_n$. Hence $L(t) = N^e(t) - N^l(t)$. Let be (if the limit exists)

$$Th := \lim_{t \to \infty} \frac{N^e(t)}{t}$$

the arrival rate into the system,

$$avgCT := \lim_{n \to \infty} \big(\frac{1}{n} \cdot \sum_{j=1}^{n} CT_j\big)$$

the average cycle time the unit spends in the system,

$$avgWIP := \lim_{t \to \infty} \big(\frac{1}{t} \cdot \int_0^t L(s)\mathrm{d}s\big)$$

the average number of units in the system.

Theorem IV.4 (Little's theorem) *If both the arrival rate Th and the average cycle time $avgCT$ exist and are finite, then the above limit in the definition of the average inventory $avgWIP$ exists and it holds:*

$$avgWIP = avgCT \cdot Th. \tag{3}$$

Corollary IV.5 *If both Th and $avgCT$ exist and are finite, then the departure rate exists and equals the arrival rate:*

$$\lim_{t\to\infty} \frac{N^l(t)}{t} = Th.$$

Little used a stochastic framework to define and prove of what is known as Little's Law, the approach we are presenting makes no stochastic assumptions, i.e., the quantities and processes are deterministic. There are other versions of Little's Theorem that allow batch arrivals, see section 6.2 of [17].

D. Calculation of the Flow Factor

Next, we establish a formula for the calculation of the flow factor for the processing line. For this, we restrict to the following queueing model: The *adapted queueing model* is based on the one given in Subsection IV-A with the following modifications:

- Units can enter and leave the system only through a finite number of gates.
- Each gate on the entering side has its correspondence on the exit side.
- The entering and the corresponding exit gate are connected by a lane.
- Once, the person entered the system, he can move forward only on the lane set up by the entering gate. He cannot switch the lane or leave the system except the exit gates.
- Each lane contains a number of clerks, not defined in detail, such that before each clerk an internal queue is formed and the clerk does not necessarily process the requests instantly.
- The sum of the time the clerks process the requests of a person during his/her voyage through a given lane (i.e., the raw process time) does not depend on the particular person involved. Hence, the system has a predefined raw process time (RPT^l) for each lane, i.e., the sum of the time the clerks process the requests of a person during his/her voyage through the lane.

Table I illustrates the queueing model.

Table I. ILLUSTRATION OF A QUEUEING SYSTEM WITH 5 LANES $l_1, l_2, \ldots, l_5$ AND MAXIMAL 8 PROCESSING STEPS AT EACH LANE.

l_1	$\Rightarrow$	■	■		■			■	■	$\Rightarrow$
l_2	$\Rightarrow$	■		■	■	■	■		■	$\Rightarrow$
l_3	$\Rightarrow$	■	■				■	■	■	$\Rightarrow$
l_4	$\Rightarrow$	■		■		■			■	$\Rightarrow$
l_5	$\Rightarrow$	■					■		■	$\Rightarrow$

We will denote by L the set of the lanes and by Th^l the throughput at lane $l \in L$.

Definition IV.6 (Flow factor) *Let $\{u_1, u_2, u_3, \ldots\}$ be the ordered list of units which enter the system, such that u_i enters the system at time t_i and $i < j \Rightarrow t_i \leq t_j$. The cycle time CT_i a unit $u_i \in U$ spent in the system can be split into the waiting time (WT_i) and raw process time RPT_i such that $CT_i = WT_i + RPT_i$. If the limit exists, then the (average) flow factor $avgFF$ is defined as:*

$$avgFF := \lim_{n\to\infty} \frac{\sum_{i=1}^{n} CT_i}{\sum_{i=1}^{n} RPT_i}. \tag{4}$$

Remark IV.7 *If $CT := \lim_{n\to\infty}\left(\frac{1}{n} \cdot \sum_{i=1}^{n} CT_i\right)$ and $RPT := \lim_{n\to\infty}\left(\frac{1}{n} \cdot \sum_{i=1}^{n} RPT_i\right)$ exists (and are finite), then the above limit exists and it holds:*

$$avgFF = \frac{CT}{RPT}.$$

Corollary IV.8 (Representation of raw process time) *Let $N_l^e(t)$ be the number of units which entered the lane $l \in L$ until time $t \in T$. Let $n := N^e(t_n) := \sum_{l\in L} N_l^e(t_n)$. Then, the average raw process time RPT of the whole processing line can be represented as follows:*

$$RPT := \lim_{n\to\infty}\left(\frac{1}{n} \cdot \sum_{i=1}^{n} RPT_i\right) = \sum_{l\in L} \frac{Th^l}{Th} \cdot RPT^l. \tag{5}$$

Hint *We obtain:*

$$\frac{1}{n} \cdot \sum_{i=1}^{n} RPT_i = \frac{1}{n} \cdot \sum_{l\in L} \sum_{i_l=1}^{n} RPT_i^l = \sum_{l\in L} \frac{N_l^e(t_n)}{N^e(t_n)} \cdot RPT^l.$$

Since

$$\lim_{n\to\infty} \frac{N_l^e(t_n)}{t_n} \cdot \frac{t_n}{N^e(t_n)} = \frac{Th^l}{Th} \;\forall l \in L$$

it follows that

$$RPT := \lim_{n\to\infty}\left(\frac{1}{n} \cdot \sum_{i=1}^{n} RPT_i\right) = \sum_{l\in L} \frac{Th^l}{Th} \cdot RPT^l.$$

■

Corollary IV.10 (Representation of flow factor) *Assumed that the conditions of Little's Theorem are satisfied. Then, the flow factor can be represented as follows:*

$$\frac{1}{avgFF} = \sum_{l\in L} \frac{WIP^l}{WIP} \cdot \frac{1}{FF^l}. \tag{6}$$

Hint *Easy calculations using Little's Theorem for each lane $l \in L$ yields to the relationship between the flow factor for the whole system and the flow factors of its components / lanes as given in* (6). ■

We can calculate the average number of units in the system first by considering the whole system and secondly considering the reduced system with one lane. Little's formula is valid in both cases. Since units cannot switch to another lane, it follows that

$$WIP = \sum_{l\in L} WIP^l.$$

Using Little's formula and the definition of the average cycle time it follows that:

$$\lim_{n\to\infty}\left(\frac{1}{n}\cdot\sum_{i=1}^{n} CT_i\right) = CT = \sum_{l\in L}\frac{Th^l}{Th}\cdot CT^l.$$

Hence, as expected:

$$avgFF = \frac{\sum\limits_{l\in L}\frac{Th^l}{Th}\cdot CT^l}{\sum\limits_{l\in L}\frac{Th^l}{Th}\cdot RPT^l} = \frac{CT}{RPT}.$$

Let us suppose that the service center has different departments, such as for "incidents", "problems", "projects", "releases", etc., which operate independently. By abstracting those departments as lanes and calculating for each department the flow factor, the flow factor of the service center can be established as in (6).

Moreover, Equation (6) determines the correlation between the flow factors of each department and the flow factor of the data center. Thus, the flow factor of the data center can be improved within an existing budget, for example, by resource reallocation, if the flow factor of some departments will be improved and the flow factor of some other departments will be degraded, see also the discussion regarding the operating curve.

A formula of the type given in (6) was proposed by Hilsenbeck in [20, p. 36]:

$$avgFF = \sum_{l\in L}\frac{Th^l}{Th}\cdot FF^l. \tag{7}$$

It seems that the formula (7) is empirical. In particular, no proof of the formula was given.

The flow factor plays an important role in the operating curve management. The operation curve follows from Kingman's equation [21]. One of the representation of the operating curve is based on the following formula (see [22, pp. 55, 58], [20, pp. 41, 44], [23] [24]).

$$avgFF = f(U) := \alpha\cdot\frac{U}{1-U} + 1. \tag{8}$$

U is the *utilization*, i.e., the percentage of the capacity $Capa$ of a tool or production segment (see [25] for a definition and [22, p. 57] for calculation). Introducing $avgFF$ and U in (8), the value for the coefficient α (variability) follows.

The operating curve as a function $avgFF(U)$ can be drawn. However, this relation is rather abstract. Since it holds

$$U = \frac{Th}{Capa}, \tag{9}$$

the flow factor in terms of a function $avgFF = f(U)$ can be easily transformed into a function of the type $CT = g(Th)$ (see [22, p. 40]):

$$CT = g(Th) := \alpha\cdot\frac{Th}{Capa - Th}\cdot RPT + RPT. \tag{10}$$

This relation is more practical as it shows how the throughput directly influences the cycle time. The self-generated graph of the function g is depicted in Figure 1. It is assumed that the average minimal cycle time RPT is 1 hour and that the maximal capacity is $Capa = 1000$. If Th is close to $Capa$, then the graph of g grows asymptotically. Hence, a point at the graph (named *operating point)* has to be chosen, such that a minimal increase of the number of items does not lead to dramatically increased cycle time. The operating curve has been used by Qimonda to improve overall fab performance.

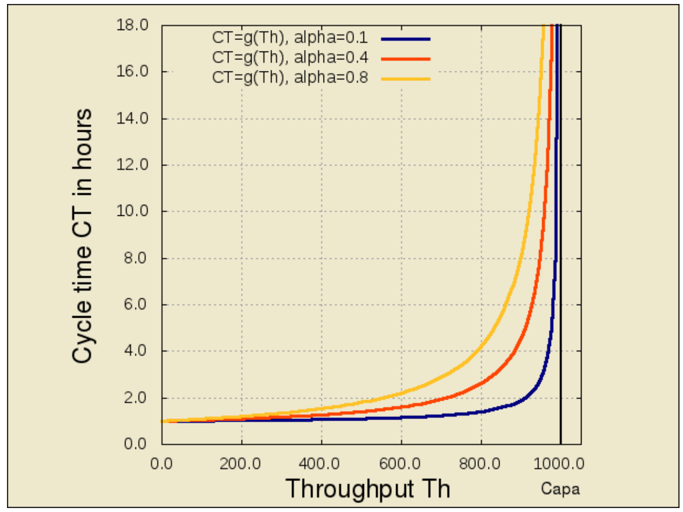


Figure 1. Graph of function g given in (10) (Operating curve). Throughput Th versus cycle time CT for four different values of α.

E. Examples

In the following, we consider three simple examples in order to illustrate the methodology used to determine Little's relation and the Flow Factor for the line.

1) Process flow considering a single step: To start with, we consider an example with one single step as shown in Fig. 2. On the rectangular scheme, the units are placed horizontally, the columns are the times at which the state of a unit can change.

Our system abstracts the process flow at an arbitrary step s. The system is in stable state, i.e., the process flow is repeated every 24 hours. Hence, we consider an arbitrary day D. As illustrated in Fig. 2 the unit u_1 enters the system at 12:00 on the previous day of day D, waits till 0:00 (pictured using gray boxes), and it is processed from 0:00 till 4:00 (pictured using a black box) when it leaves the system. For $i \in \{1,2,3,4\}$ the unit u_{i+5} has the same behavior as the unit u_i but it is shifted by 24 hours. It is quite easy to see that the number of units that left the system on day D, denoted by Th, is equal to five units per day.

We use the relation in (1) to calculate the average WIP with the hour as the lowest granularity for the time, hence, the cardinality $|D|$ of D is 24. Between 0:00 and 4:00, there are two units in the system, between 4:00 and 8:00, there are three units and so on. Thus, $avgWIP_s^D = (1+2+3+4+4+3+2)\cdot 4/24 = 19/6$ units.

In order to calculate $avgCT$ we observe by counting the cubes that u_1 spent $16/24$ days in the system. The units u_6 and u_7 are not considered, since they did not leave the system on day D. Hence, the average cycle time which a unit spent in

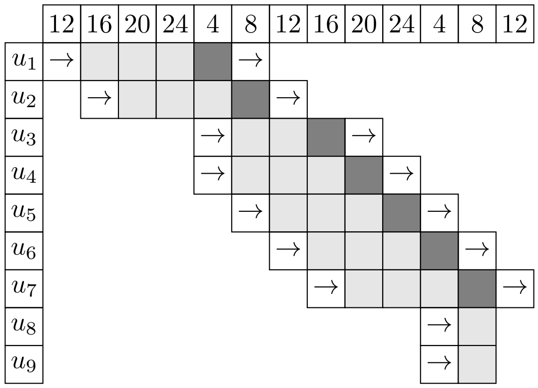


Figure 2. Process flow at an arbitrary step. The rows of the rectangular scheme represent the line of the system, the columns are the times at which the state of a unit can change. Arrows indicate when a unit is entering or leaving the system. Wait states of units are depicted using light gray boxes, corresponding process states are depicted using dark gray boxes.

the system is equal to $4 \cdot 19/(5 \cdot 24)$. Obviously, the relation $avgWIP = avgCT \cdot Th$ as given in (2) of Lemma IV.1 holds.

2) Process flow considering two steps: Subsequently, in order to illustrate the process flow of the line, we consider a more complex example with two steps as shown in Fig. 3. The first day is not part of the stable phase, it just shows the ramp up phase. As in the previous example u_6 is the follow up of u_1, u_7 is the follow up of u_2, etc.

For example, the unit u_2 enters the system at 4:00 on the first day, it is processed from 8:00 to 12:00 at the first step (dark gray box), then waits till 0:00 next day (gray box), it is processed between 0:00 and 4:00 (black box) and leaves the system. Analogously, the unit u_6 enters the system the first day at 16:00, waits at the first step till 4:00, then it is processed for 4 hours at the first step (dark gray box) and remains in waiting position (gray box). The unit u_{10} enters the system the second day, but it is not processed on that day. Regarding the second step, we have $avgWIP = 22/6$ since u_1 is not anymore in the system for the considered day and $avgCT = 4 \cdot 22/(25 \cdot 5)$.

For the whole system, we can calculate $avgWIP$ by considering the second day (when all the units from the two operations are in the system: $avgWIP = 41/6$ and $avgCT = 41/30$. Please be aware that u_2 spent additional $2 \cdot 4$ hours in the system due to waiting from the previous day. One can observe this by looking at u_7, which behaves like u_2, but only postponed by one day. Obviously, Little's formula holds. As expected, the average cycle time for the whole system is equal to the sum of the cycle times for the individual operations.

3) Process flow considering two departments: Finally, for the calculation of the flow factor FF_{line} for an entire line, we consider an example as illustrated in Fig. 4. In our example, the line consists only of one step and has two departments, d_1 with units $\{u_1, u_2\}$ and d_2 with units $\{u_3, u_4, u_5\}$. The units u_6 and u_7 are the follow-up units for u_1 and u_2, respectively.

The flow factor FF_{d_1} is equal to the total time the units of product d_1 spent in the system divided by the time the units were processed, i.e., $FF_{d_1} = (2+2+4+2)/(2+2) = 10/4$. Same considerations give $FF_{d_2} = 11/3$.

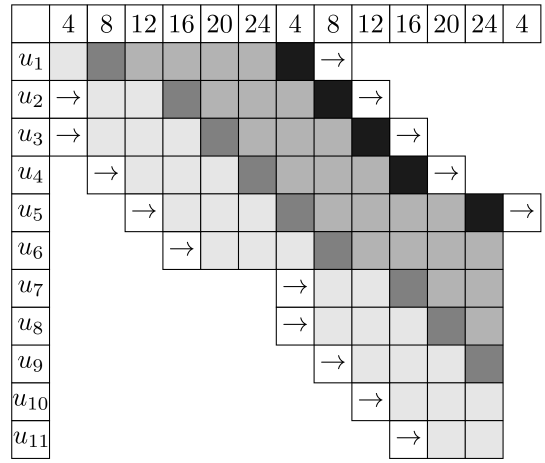


Figure 3. Process flow considering two steps. The rows of the rectangular scheme represent the line of the system, the columns are the times at which the state of a unit can change. Arrows indicate when a unit is entering or leaving the system. Wait states of units related to step 1 are depicted using light gray boxes, corresponding process states using dark gray boxes, wait states of units related to step 2 are depicted using gray boxes, corresponding process states are depicted using black boxes.

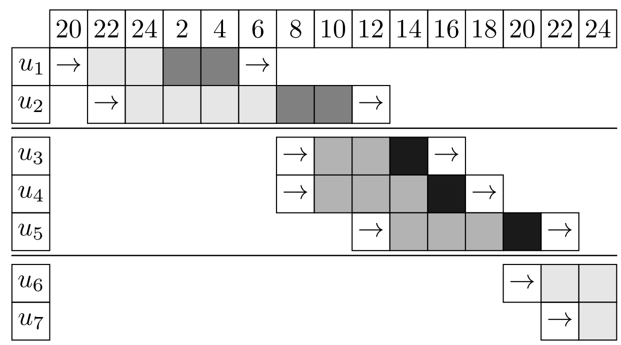


Figure 4. Process flow considering two departments, each containing one step. The units $\{u_1, u_2\}$ belong to department d_1, the units $\{u_3, u_4, u_5\}$ belong to department d_2. The units u_6 and u_7 are the follow-up units for u_1 and u_2, respectively. The rows of the rectangular scheme represent the lines of the system, the columns are the times at which the state of a unit can change. Arrows indicate when a unit is entering or leaving the system. Wait states of units belonging to department 1 are depicted using light gray boxes, corresponding process states using dark gray boxes, wait states of units belonging to department 2 are depicted using gray boxes, corresponding process states are depicted using black boxes.

The flow factor for the entire line FF_{line} can be calculated analogously as the total time the units spent in the system divided by the time the units were processed, i.e., $FF_{line} = (10+11)/(4+7) = 22/7$. Same considerations as in the previous example (counting the boxes) give $avgWIP_{d_1} = 10/12$ and $avgWIP_{d_1} = 11/12$. Hence $avgWIP_{line} = 21/12$. Obviously, formula (6) for calculating the flow factor of a line by considering the flow factor of its components holds.

V. Additional Strategies

A. Continuous Change Strategy

We will present the advantages and the disadvantages of the Continuous Change Strategy (CCS) by means of a simple example. The CCS represents an enhancement of the classical release strategy: A complex major release is split into a certain number of minor releases, such that by performing all the minor releases, the major release is accomplished. Thus, instead of a big jump, some small steps are taken toward the goal. The CCS is not reduced to the release strategy, it can be applied to all fields of change management, such as quality improvement, project management, etc.

The approach of the CCS is closely related to Scrum [26], since releases can be regarded as software projects. The approach of Scrum is based on the experience that many development projects are too complex to be captured in a fixed plan. The long-term plan is refined and improved continuously, detailed plans are conceived only for the next development cycle.

This way, we assure that [27]:

- we keep as many options open as possible,
- we accept that it is not possible to do things right from the beginning,
- irrespective of the starting point, we realize that it is important to learn fast enough from one's own failures, feedbacks and achievements,
- we favor an adaptive, investigative approach to a rigid, planed concept.

More formally, let be S_k the state of the system before the major release R and let S_{k+1} be the state of the system after implementing the major release R. Let $r_{i_1}, r_{i_2}, \ldots, r_{i_n}$ be the succession of the minor releases such that r_i transforms system S_i into system S_{i+1}. Hence, the sequence $S_{i_1}, S_{i_2}, S_{i_3}, \ldots, S_{i_n}$ such that $S_k \equiv S_{i_1}$ and $S_{i_n} \equiv S_{k+1}$ represents the evolution of the system during the minor release implementation. The state S_{k+1} corresponding to upgrade to the major release, is achieved after implementing the last minor release S_{i_n}. All the states S_i with $i \in \{i_1, i_2, \ldots, i_n\}$ are stable states, such that productive service can be provided.

The major professional challenge of this strategy lies in the difficulties to design a step by step upgrade strategy. In some circumstances, this might be a very sophisticated and time consuming endeavor. However, the complexity of a one step release change is also not negligible.

The lessons learned by the first author during his long term project experience at a semiconductor company were that a multi-choice strategy is crucial for the success of complex upgrades. This way, if a continuation from one point was not any more possible by reaching a deadlock situation, a fall back to a previous step can be considered, which bypasses the deadlock.

As in the example of Figure 5, the evolution of the minor releases S_{i_1} to S_{i_2} is linear, but the upgrade to the release S_{i_3} fails. Due to deadlock property at release S_{i_3}, a fall back to S_{i_2} is necessary. Accordingly, to avoid an impasse, both an upgrade and a downgrade strategy have to be developed, the evolution of the sequence of the minor releases has to be adjusted accordingly. A new – ad hoc developed – series of minor releases $S'_{i_3}, S'_{i_4}, \ldots, S_{i_n}$ is used instead of the initial $S_{i_3}, S_{i_4}, \ldots$ ones. Due to the complex and unpredictability behavior of the minor releases, a detailed specification of

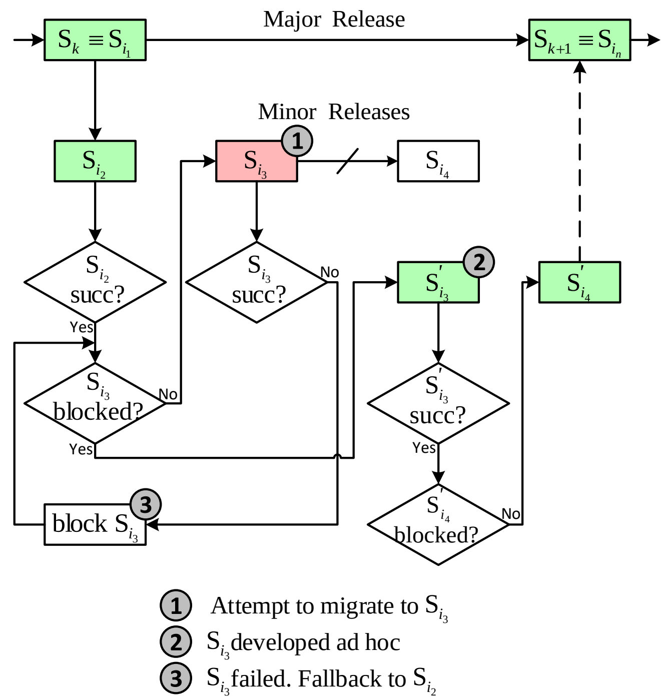


Figure 5. Example of the evolution of the system during minor releases. The actual sequence of the successful minor releases is depicted in green.

the migration strategy is meaningful only for the succeeding release. Less detailed, but flexible migration plans have to be conceived for the successive further releases. If the migration to an appropriate minor release fails, the whole subsequent design of the minor releases has to be reconsidered.

The ability to effect change continuously has become an increasingly necessary core competency [28]. Change should not be a turbulent, anxiety-inducing event, but a part of the everyday routine. There is a natural opposition to change, people are reluctant to change for various motifs and considerations, some of them are not transparent or comprehensible by rational evidence. Some of the people fear for their job, others just do not have time to become involved in an overall change. Thus, departments implementing CCS can avoid big fluctuation in manpower, can much better control costs and avoid the risks due to an overall change. Some of the drawbacks of CCS are:

- The contiguous need of highly qualified people, with skills beyond the maintenance task,
- the inflation of some department with personnel, which imply higher logistics,
- increased responsibility for the change, if the alternative overall change was outsourced,
- increased effort to work out the strategy on how to split the overall change into small steps.

The duality of continuity and change has conventionally been treated as a dilemma, i.e., either to be in continuity or to change [29] [30]. Accordingly, we have chosen the term *continuous change* to be in accordance with the present day terminology and thus to avoid discussions which are out of our scope. We mean by term *continuous change* a change involving smooth, seamless, and non-disruptive intermediary steps as stated above.

Studies show that on company level the failure rate for organizational change projects has stayed constant from the 1970' to 2013 at 60 to 70% [31]. For example, regarding the significant IT-projects at Qimonda, the push through rates was not very high, regarding the less important IT-projects it was even much lower. The major cause for the failures was the underestimation of the technical and logistical difficulties and a poor preparatory phase. One of the measures we took after one of our most important strategic project failed, was to limit the maximum duration allowed for the IT-projects to six month, whereby each project had its own justification, a decision towards CCS.

A similar and/or pursuing concept is the Continuous Delivery, it comprises reliable software releases through build, test, and deployment automation [32] [33]. Humbles's [33] advice: *If you do nothing else, start automating your deployments*.

Further thoughts and development on this area could include studies concerning the return on investment by switching from a classical release strategy to CCS. Redefining on the fly the goals of the major upgrade projects could increase flexibility and adaptability to the new challenges of the company.

B. Customer Dissatisfaction

We are not aware of any reliable method to accurately identify customer satisfaction (CS). The usual method of questionnaires does not deliver credible outcomes, since:

- the survey is a posteriori and the customer has no possibility to actively influence the outcome of the results and hence he is not really interested in the conclusions of the survey,
- the customers who take part in the survey may not be representative,
- the customer may not be convinced that his/her proposals will be implemented, then why bother?

In order to circumvent this impediment, we will focus on the analysis of customer dissatisfaction (CD). Although it seems that CD is the opposite of CS, this is only valid with limitations due to the way we determine them.

There is no possibility to measure CD or CS in a straightforward way – like for example, for the cycle time – and associate a number to it, which is meaningful per se. Values of dissatisfaction can be *neutral, low, medium, high*, etc. Irrespective the impossibility to measure the degree of dissatisfaction as mentioned above, conclusive metrics can be set up and used to measure the variation of the CD in a automated way, such that it is independent of the customer‘s perception of good or bad.

Let us consider for example, the ticket system. The classical way of determining CS is through a questionnaire. Although this strategy seems to deliver satisfactory results of the questionnaires are very well prepared and the people taking in the survey are representative, the results can be altered by using a successful public relation. Accordingly, just by making the impression to take care of the customers, without really improving the QoS, the results of the questionnaire can be substantially improved. The drawback of this methodology, is that on one side, the results are so good that no improvement measures are taken, on the other side, the customers realize that the promises were not kept and the loose confidence.

We can circumvent this anomaly by considering the dissatisfaction as our main goal to be determined. In order to be able to do this, some adaptations of the work flow are necessary. Accordingly, each ticket has some attributes, like *priority, weight, escalation, remote access, technical skills*, etc. The *priority* represents the speed (time) of the ticket through the processing phase. The *weight* of the ticket represents the importance of the resolution of the ticket for the production process. The *escalation* represents the explicit discontent with the progress of the processing of the ticket, whatever the reasons are. The *remote access* represents the wish of the customer to be assisted by remote access. The *technical skills* represents the level of explanation of the solution towards the customer, such that technically skilled customers will get a less detailed explanation of the solution and vice versa. Of course, other attributes can be meaningful to consider.

The general assumption is that the customer will react in order to improve the processing of his/her ticket if it does not fulfill his/her expectation. This assumption may seen reasonable as long as the customer hopes that through his/her steps taken he can influence the outcome of the solution of the ticket. If the customer loses confidence in the system, then the above assumption does not hold.

The strategy to track the dissatisfaction is straightforward, each ticket is started with its attributes set to the minimal values as default. Is the customer satisfied with the default priority, which corresponds for example, to one week processing time of his/her ticket, then he will take no action at all. Is he not satisfied, then he will have to ask for an increase of priority by answering some questions regarding the need for increased velocity of his/her ticket. By making this extra effort and considering the time involved as a promising investment, the customer gives important hints regarding his/her expectations regarding the processing of the ticket. As mentioned above, the priority of the ticket should be increased as the result of the attempt or alternatively, the decision should at least be sufficiently justified.

Valuable information can be obtained by an ingenious system of measures and metrics. For example, an increased number of people who requested a remote access is not a sign for a bad documentation per se, but if this number increases dramatically for highly skilled customers than this is clearly a sign for outdated and poor documentation and for increased customer dissatisfaction. Analogously, an increased number of responses from the help desk side for highly weighted tickets points to the incapacity of the help desk to resolve at least the critical tickets in a satisfactory manner.

These objective methods provide a trend analysis regarding CD, if the results are representative enough. However, these methods can be combined with questionnaires. Each time an attribute of a ticket is changed, the customer has to complete a small questionnaire. However, this way the opinions of those who are discontent, may prevail. Additionally, when closing the ticket, a survey can be conducted.

Through the investigations on customer dissatisfaction a paradigm change from a predominantly subjective methodology – like the questionnaires – to determine the service quality, towards objective, measurable procedures can be initiated.

Customer satisfaction is one of the most important concepts in economic research literature, having been the focus of countless studies. When compared with the literature on satisfaction, the concept of dissatisfaction has been the focus of a far fewer numbers of studies [34]. Most of the researchers have seen dissatisfaction as either the opposite of satisfaction or as of different significance, for additional details see [34].

CS implies exceeding the customer's expectations and it is associated with positive feelings, whereas CD is more or less an affective reaction associated with negative emotions, which are more intense and remain much longer in memory than positive ones, for a broader treatise see [34] [35].

To summarize, CD is seen by researchers as a behavioral response to failed service encounters [35], whereas within our approach, CD is an impersonal, objective state of the customer on his/her way to achieve his/her objectives. By following a dispassionate, factual customer dissatisfaction strategy, we pave the way for measuring the dissatisfaction in an objective manner, thus being able to set up metrics and indicators which are independent of the disappointments of the individual customer.

Our objective approach as above, can only be applied if the customer is able to actively influence the QoS through his/her behavior. For example, the degree of dissatisfaction of a customer who lost his/her checked-in baggage is unpredictable [36]. Certainly, an attentive reaction and customer-friendly measures of the airline company can substantially reduce dissatisfaction or even to lead to appreciation.

Regardless of how dissatisfaction is determined, by lowering the barriers to complaining, the percentage of customers who articulate their problems can be increased effectively [37]. In addition, the customers who do not believe that the management will take appropriate measures, have also to be reached.

Captive services are services which are provided in systems without competition, either directly or through a process which limits the consumer's choice, its control and power; service captivity is a consumer's perception that s/he has no options for obtaining a service other than the current provider [38]. Some of the services provided by data centers are of this nature, customers cannot choose another data center without leaving the company to which the data center belongs.

The results of the inquiry [39] shows that the captive services have their own characteristics, which need an exclusive promotion and management of the relation with clients. The captivity, in which the consumers find themselves, induce negative emotions compared to the competitive situation. The perception of QoS is therefore diminished, thus exacerbating directly or indirectly the dissatisfaction of the consumers and the negative verbal publicity. Showing more consideration in the relations with the clients and taking more account of their interests, should ensure the reduction of the negative emotions. Alternatively or additionally, by giving back power and control to the captive customers, for example, by allowing them more flexibility to choose between alternatives, could ease the monopoly situation, thus reducing their discontent and as a consequence, improve the image of the company.

VI. Use case: An excerpt

We illustrate the principles of improving the QoS of a data center by means of a simplified example. Let us consider the department which provides the e-mail service of a data center. Firstly, we establish the conditions such that the providing the service is at all possible. Secondly, we set up metrics and compose them in order to be able to track the evolution of QoS.

One of the most sensible indicators whose value has to be estimated is the raw process time RPT which is the (average) minimal cycle time to process an incident. It contains only the effective time to process the incident, for example, not including coffee breaks, private telephone calls, etc. Let us suppose that RPT is equal to 1 hour. In real systems (see [22, pp. 46, 48]) the cycle time CT corresponding to a specific throughput, denoted by Th is measured. Let us suppose that by considering the raw process time the maximum capacity $Capa$ is 1000 incidents per month.

Introducing CT and Th in (10), the value 0.4 for the coefficient α (variability) follows. As shown in Figure 1 we can easily follow that a slightly increase of the throughput (after leaving the linear part of the graph) considerably increase the cycle time. In order to avoid the flooding of the departments with tickets, the natural reaction of the employees is to reduce the raw process time and consequently reduce the QoS of the department. Hence, in our example, if the throughput exceeds 800 incidents per month appropriate measures should be taken in order to avoid the collapse of the service. On the contrary, if the throughput is equal to 400 tickets per month (being on the linear part of the graph), a part of the staff can be relocated to assist other services. The relation (6) shows the correlation between the flow factor of the individual departments and the data center and can be used to balance the individual departments.

In order to establish normalized / composite metrics, we consider those presented in Subsection III-A.

We describe below some of the metrics used in incident management, normalized and directed as described in Subsection III-B, i.e., each metric takes values in the closed interval $[0, 1]$ and a greater value for the metric implies a better accomplishment of the business requirements.

$$m_{01} := \frac{1}{\text{“Total No. of incidents”}}$$

$$m_{02} := 1 - \frac{\text{“No. of repeated incidents”}}{\text{“Total No. of incidents”}}$$

$$m_{03} := 1 - \frac{\text{“No. of repeated incidents with known solution”}}{\text{“No. of repeated incidents”}}$$

Unfortunately, the “Maximum No. of incidents” is not a priori known. Hence, generally speaking, it cannot be used in the formula. Furthermore, the business requires that corresponding measures are taken, such that repeated incidents are avoided. Therefore, “No. of repeated incidents” should be kept low. Further metrics, which are considered (SLA refers to Service Level Agreement):

$$m_{04} := \frac{\text{“No. of escalated incidents”}}{\text{Total No. of incidents”}}$$

$$m_{05} := \frac{1}{\text{“Average cycle time to resolve the incident”}}$$

$$m_{06} := \frac{1}{\text{“Average waiting time from user side”}}$$

$$m_{07} := \frac{1}{\text{“Average working time on the incident”}}$$

$$m_{08} := \frac{\text{“Total No. of incidents resolved within SLA time”}}{\text{“Total No. of SLA relevant incidents”}}$$

$$m_{09} := \frac{1}{\text{“First reaction time to repair the incident”}}$$

$$m_{10} := \frac{1}{\text{“No. of responses from service center side”}}$$

Unfortunately, defining metrics fulfilling the conditions as above, is not always straightforward. Let us consider $Incid_{out}$ as the total number of incidents closed and $Incid_{in}$ as the total number of incidents opened in the time frame considered. In order to avoid the flooding of the data center with incidents the metric $k := Incid_{out}/Incid_{in}$ could be tracked. Unfortunately, this metric does not fulfill our requirements, since it can take values outside the interval $[0, 1]$. In order to avoid this impediment, we define $k_{in} := Incid_{in}/$“Total No. of incidents” and $k_{out} := Incid_{out}/$“Total No. of incidents” and set

$$m_{11} := \frac{1 + k_{out} - k_{in}}{2}.$$

Then, m_{11} is normalized and satisfies the above conditions imposed for metrics. Generally speaking, the effective minimal and maximal value of a metric is not known, nor is the distribution a priori known. Thus, for example, a metric m_1 takes values in the interval $[0.5, 0.6]$ and another metric m_2 in the interval $[0.2, 0.7]$ with nearly uniform distribution. Hence, the metric m_2 varies more widely than m_1 and this should be considered – for example, using the standard deviation – when setting up the groupings for the composition of the metrics, such that metrics having a low standard deviation should be assigned to more important groups.

Now, let us consider in our use case five composition groups, $G_0, G_1, \ldots, G_4$, such that G_1 is the most relevant group. Let us associate the weight w_i to group G_i and let us consider the weighting according to an exponential function, such that $k_0 := 1$ and $k_i := w_{i-1}/w_i$ for $i > 0$. This yields to the relation: $w_i = w_0/\prod_{l=0}^{l=i} k_l$. In our example $k_1 := e^1$, $k_2 := e^{0.75}$, $k_3 := e^{0.5}$, etc. The values for k_i are illustrated in Figure 6. Let us assign the above metrics to the composition groups, such that index set of the metrics assigned to the group G_l is equal to I_l and let n_l the number of metrics in the group G_l. Then, according to the composition rules: $\sum_{i=0}^{4} n_i \cdot w_i = 1$. Hence, the weight values follow. The value of the composition metric M is:

$$V(M) := \sum_{i=0}^{4} w_i \cdot \sum_{l \in I_i} V(m_l).$$

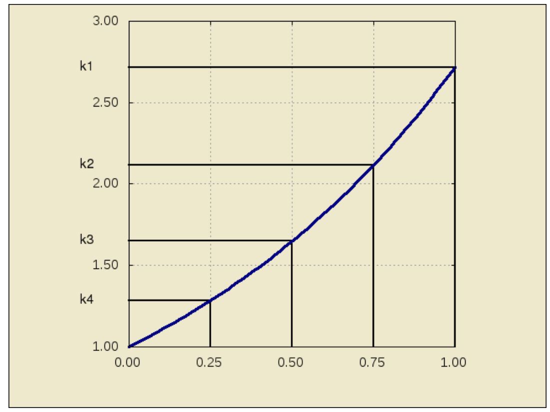


Figure 6. Graphical construction of values k_i such that metrics are weighted according to an exponential function, depicted for $f(x) = e^x$.

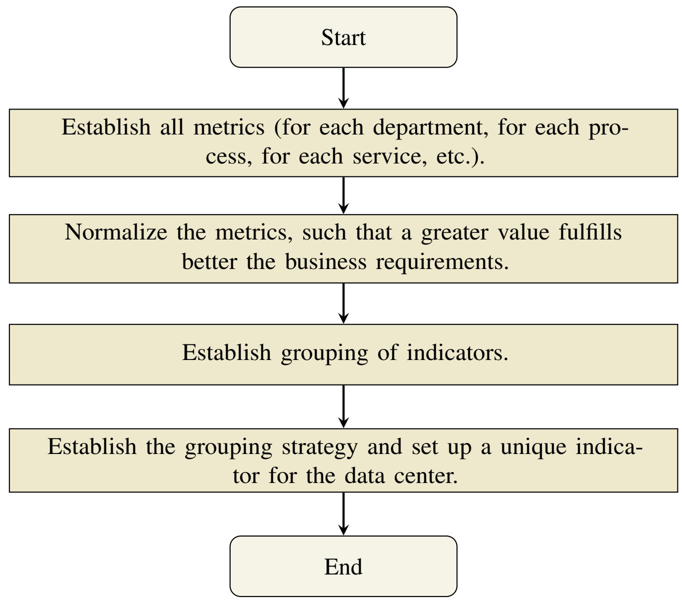


Figure 7. Simplified flow diagram regarding the composition strategy.

Examples of services at the ZIH of the Technische Universität Dresden (TU Dresden) are: “E-Mail Service”, “Backup and Archive Service”, “Data Exchange Service”, “Access to High Performance Computing Resources”, etc. [40] We conclude this section by presenting the assistance system for a data center in Figure 8 and by summarizing the composition strategy via the flow diagram given in Figure 7.

Next, we provide a simplified example of an incident management process with special emphasis on the customer dissatisfaction strategy. This process is depicted in Figure 9. Each request to the help desk is captured by the ticket system. At creation time, each ticket comprises attributes (like priority, weight, escalation, remote access, technical skills, etc.), which are set to the most preferred predefined values. For example,

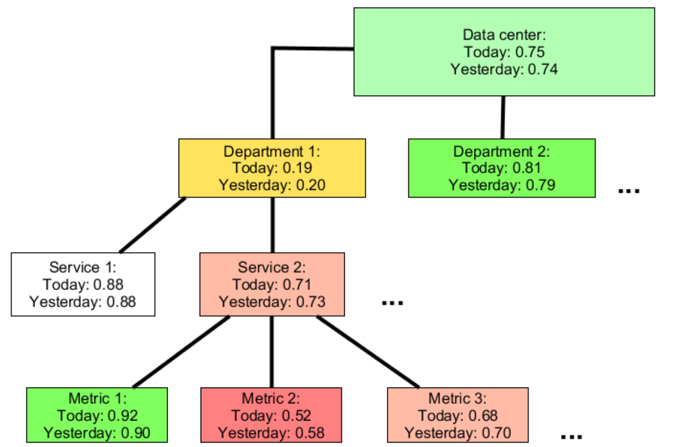


Figure 8. Hierarchical assistance system for a data center consisting of different departments, services, and metrics as levels. Changes will be depicted in green, white or red depending on the indicator's values of today and yesterday.

priority, weight is set at its lowest value, remote access to "No", and technical skills are set at their highest rank, i.e., at the least detailed explanation. The customer can change these default values any time during the life time of the ticket, provided the ticket is ready for processing at the customer side.

The ticket is initiated by the customer and then submitted to the help desk. The ticket is analyzed at the help desk, the ticket's attributes are updated depending on the capabilities of the help desk. If the help desk can solve the problem by themselves, then the solution – corresponding to the accuracy specified through the technical skills – is forwarded to the customer.

If the help desk does not have the expertise to solve the problem, the ticket is forwarded to the domain experts for further processing. The solution in this case is sent back to the help desk, which redirects the ticket to the customer. If remote access is necessary, a skilled specialist will get in contact with the customer to enable remote connection and on-site solution.

If the customer is satisfied with the provided solution, then he closes the ticket on his/her side, else he forwards the unresolved part of the problem to the help desk and the procedure loops until a satisfactory solution is found or the problem turns out to be unsolvable. In both cases, the customer or/and the help desk closes the ticket.

One of the major tasks of the help desk is documenting the findings accurately. By evaluating the feedback of the customers regarding ticket attributes like priority, weight, escalation, remote access, etc., the help desk can obtain valuable information regarding the degree of dissatisfaction with the service it provided.

VII. Conclusion and Future Work

The basic research idea regarding the QoS was to establish objective criteria in order to determine the goodness of services, such that this goodness should be unequivocally measured and therefore be expressed numerically. Therefore, the goodness of services established as above would be independent of the personal assessment of beneficiaries of the service.

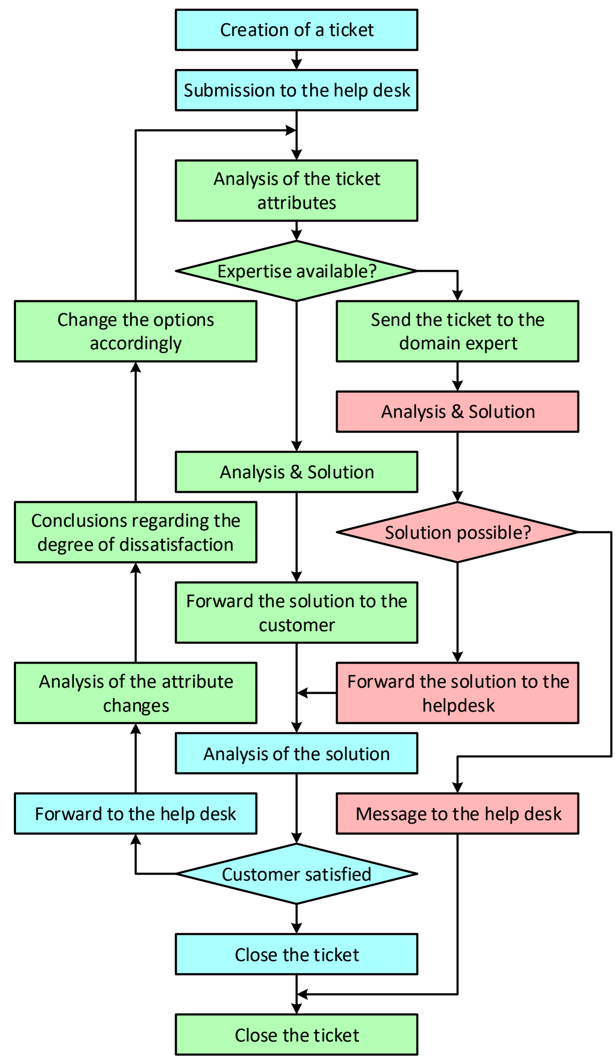


Figure 9. Simplified example of an incident management process with special emphasis on the customer dissatisfaction strategy.

The classical approach to estimate the QoS is through questionnaires and interviews. This approach constitute an extremely valuable source of information. However, it comprises inadvertently the degree of expectation towards the quality of the service offering. The main benefit of our proposed method against the classical approach is a reproducible, straightforward method. It defines the quality in relative terms and should visualize the increase or decrease of the QoS pretty accurately.

The presented methodologies to increase QoS, such as the Composition of Metrics (CoM), the Queueing Theory including the Composition of the Flow Factor (FF), the Continuous Change Strategy (CCS), or the Customer Dissatisfaction (CD), are not interrelated and can be used independently of each other in accordance with the priorities / benefits one would like to achieve or the costs that are taken into consideration. In brief summary, CoM combines various metrics to a new synthesized compound metric. The fluctuation of these metrics reproduces the quality variation of the respective service. The application of the Queueing Theory and FF avoids congestion and bottlenecks. CCS assures merely a smooth transition during version change. Some of the above methodologies yield similar results, other methodologies pursue different objectives and are therefore complementary. Hence, there is no road map or prioritized strategy to follow. For example, if a version change is very accurately prepared, then the number of tickets due to the version change will remain within acceptable limits. This way, the impact on the workload of the help desk of the service provider will not increase substantially during the transition period. On the contrary, if there is sufficient reserve regarding the workload at the help desk, then a less carefully prepared version change will not have dramatic effect on the proper functionality of the help desk. In this latter case, the staff of the help desk will have enough time to deal with the increased number of tickets.

The main conclusion of this paper is that there exists a positive answer to the basic research idea mentioned above. Objective criteria can be set up in order to measure the QoS. Many aspects of the technologies presented in this paper in order to increase the QoS are new to our best knowledge. An example is the formula to calculate the FF of the data center out of the FF of each department as given in (6). In order to demonstrate the above formula, the corresponding theory has been formulated. The FF plays a crucial role in the operating curve management used to improve the performance of data centers. The approach of the CCS, although closely related to Scrum [26] and that of the CD, contains a unique combination of practical knowledge and experience of the authors in the field of business-critical operation of computer centers. We are not aware of any discussion regarding the strategy of synthetic compound metrics in the scientific literature.

We set up a formal, mathematical model and analyzed the QoS and the modalities to enhance it within this model. In that way, the QoS provided by the TU Dresden can be improved, which implicitly leads to a good ranking of the TU Dresden between the universities in Germany and world wide. By extending our investigation to a formal model, we can conclude that our results remain valid in application to other domains as long as the new domain can be mapped to the existing mathematical model.

The key result of our research is that from a service provider perspective QoS can be characterized in mathematical terms pretty accurately. The improvement / degradation of the overall service or part of it can be tracked in IT systems and can be visualized through GUIs. According to the definition of ITU-T Rec. E.800 [41], QoS is the "Collective effect of service performances which determine the degree of satisfaction of a user of a service". The long term experience of the first author, working in the semiconductor industry is very similar to the definition above, such that it does not suffice to consider only the service provider perspective. The users perception of the QoS should also be considered. Further research is necessary to establish the correlation (or lack of it) between the objectively improved service and the subjective perception of the customer.

The composition strategy of various metrics to form an overall indicator can be a very complex endeavor. If all the metrics improve or degrade, then the overall indicator will improve or degrade accordingly. The question is in which direction will the overall QoS indicator swing if some metrics improve, some other degrade in time. We are not aware of any research in this direction. Similarly, how can the overall QoS be enhanced within the limited budget by improving some components and degrading others by resource reallocation.

The study has been accomplished for the data center of the ZIH, TU Dresden. However, it can be used to improve the QoS by any service provider in the event that the real world can be mapped to the formal model used in this approach.

The similitude between between a data center and a semiconductor fab regarding performance improvement cannot be denied. It would be then advantageous to identify the major differences, such that the theory developed to improve the performance of a semiconductor fab could be adapted for data centers. This work is a little step in this direction.

ACKNOWLEDGMENT

Part of this paper (including the study regarding Little's Theorem) completes an unfinished study of the first author regarding the performance of semiconductor fabs within the scope of the Cool Silicon Project (2012 - 2014). Furthermore, we acknowledge the assistance and helpful comments of the anonymous referees.

REFERENCES

[1] M. Zinner et al., "Measuring and Improving the Quality of Services Provided by Data Centers: a Case Study," in Proceedings of The Thirteenth International Conference on Software Engineering Advances (ICSEA 2018), L. Lavazza, R. Oberhauser, and R. Koci, Eds., 2018, pp. 61–71, IARIA Conference. [Online]. Available: https://www.thinkmind.org/index.php?view=article&articleid=icsea_2018_4_10_10051

[2] Symantec Corporation, "State of the data center survey – global results," September 2012, retrieved: May 2020. [Online]. Available: http://www.symantec.com/content/en/us/about/media/pdfs/b-state-of-data-center-survey-global-results-09_2012.en-us.pdf

[3] G. I. Butnaru, "The quality of services in tourism and in the romanian accommodation system," Analele Stiintifice ale Universitatii "Alexandru Ioan Cuza" din Iasi - Stiinte Economice, vol. 56, 2009, pp. 252–269. [Online]. Available: https://EconPapers.repec.org/RePEc:aic:journl:y:2009:v:56:p:252-269

[4] S. Jain and G. Gupta, "Measuring Service Quality: Servqual vs. Servperf Scales," vol. 29, 04 2004, pp. 25–38.

[5] C. Ennew, G. V. Reed, and M. Binks, Importance-Performance Analysis and the Measurement of Service Quality, 03 1993, vol. 27, pp. 59–70.

[6] B. Edvardsson, "Service Quality: Beyond Cognitive Assessment," vol. 15, 04 2005, pp. 127–131.

[7] A. P. Parasuraman, V. Zeithaml, and L. Berry, A Conceptual Model of Service Quality and its Implication for Future Research (SERVQUAL), 01 1985, vol. 49.

[8] K. Ishibashi, Maintaining Quality of Service Based on ITIL-Based IT Service Management, 08 2007, vol. 43, pp. 334–344.

[9] E. L. Hernandez, "Evaluation Framework for Quality of Service in Web Services: implementation in a pervasive environment," Master's thesis, INSA Lyon, France, 2010.

[10] itSMF UK, ITIL Foundation Handbook, 3rd ed. Norwich: The Stationery Office, 2012.

[11] ServiceNow, "Key Performance Indicators (KPI) Examples, Dashboard & Reporting," 2018, retrieved: May 2020. [Online]. Available: http://kpilibrary.com/

[12] D. Parmenter, Key Performance Indicators: Developing, Implementing, and Using Winning KPIs, ser. BusinessPro collection. Wiley, 2015.

[13] M. Zinner et al., "Automatic documentation of the development of numerical models for scientific applications using specific revision control," in ICSEA 2017, The Twelfth International Conference on Software Engineering Advances, L. Lavazza, R. Oberhauser, R. Koci, and S. Clyde, Eds., Oct. 2017, pp. 18–27, IARIA Conference. [Online]. Available: http://www.thinkmind.org/index.php?view=article&articleid=icsea_2017_1_30_10110

[14] ——, "Revision control and automatic documentation for the development numerical models for scientific application," International Journal on Advances in Software, vol. 11, no. 3 & 4, 2018, pp. 214–226. [Online]. Available: http://www.iariajournals.org/software/soft_v11_n34_2018_paged.pdf

[15] L. Turpin, "A note on understanding cycle time," International Journal of Production Economics, vol. 205, 2018, pp. 113 – 117. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925527318303748

[16] K. Sigman, "Notes on Little's Law," 2009, retrieved: May 2020. [Online]. Available: http://www.columbia.edu/~ks20/stochastic-I/stochastic-I-LL.pdf

[17] M. El-Taha and S. Stidham Jr, Sample-Path Analysis of Queueing Systems. Kluwer Academic Publishers, 01 1999, vol. 11.

[18] J. D. C. Little, "A Proof for the Queuing Formula $L = \lambda W$," Oper. Res., vol. 9, no. 3, Jun. 1961, pp. 383–387. [Online]. Available: http://dx.doi.org/10.1287/opre.9.3.383

[19] J. Shortle et al., Fundamentals of Queueing Theory, 5th ed., ser. Wiley Series in Probability and Statistics. Wiley, 2018.

[20] K. Hilsenbeck, "Optimierungsmodelle in der Halbleiterproduktionstechnik," Ph.D. dissertation, Technische Universität München, 2005, retrieved: May 2020. [Online]. Available: http://nbn-resolving.de/urn/resolver.pl?urn:nbn:de:bvb:91-diss20050808-1721087898

[21] M. Holweg, J. Davies, and A. D. Meyer, Process Theory: The Principles of Operations Management, ser. BusinessPro collection. Oxford Univ. Press, 2018.

[22] W. Hansch and T. Kubot, "Factory Dynamics Chapter 7," retrieved: May 2020. [Online]. Available: http://fac.ksu.edu.sa/sites/default/files/Factory%20Dynamics.pdf

[23] S. S. Aurand and P. J. Miller, "The operating curve: a method to measure and benchmark manufacturing line productivity," in 1997 IEEE/SEMI Advanced Semiconductor Manufacturing Conference and Workshop ASMC 97 Proceedings, Sep 1997, pp. 391–397.

[24] W. J. Hopp and M. L. Spearman, Factory Physics: Foundations of Manufacturing Management, Burr Ridge, IL, 2nd ed. Irwin/McGraw-Hill, 2001.

[25] D. Baur, W. Nagel, and O. Berger, "Systematics and Key Performance Indicators to Control a GaAs Volume Production," 2012, retrieved: May 2020. [Online]. Available: http://www.csmantech.org/Digests/2001/PDF/12_3_Berger.pdf

[26] H. Kniberg, Scrum and XP from the Trenches. Lulu.com, 2015.

[27] H.-J. Gergs, "Neue Herausforderungen an das Change Management," in Führen in ungewissen Zeiten. Springer, 2016, pp. 189–203.

[28] T. B. Lawrence, B. Dyck, S. Maitlis, and M. K. Mauws, "The Underlying Structure of Continuous Change," MIT Sloan Management Review, vol. 47, no. 4, 2006, p. 59.

[29] P. Sushil, "Does Continuous Change Imply Continuity?" Global Journal of Flexible Systems Management, vol. 14, 09 2013.

[30] S. Nasim and Sushil, "Revisiting Organizational Change: Exploring the Paradox of Managing Continuity and Change," Journal of Change Management, vol. 11, no. 2, 2011, pp. 185–206. [Online]. Available: https://doi.org/10.1080/14697017.2010.538854

[31] R. Ashkenas, "Change Management Needs to Change," Harvard Business Review, vol. 16, no. April, 2013.

[32] J. Fish, "The Practical Guide to Enterprise DevOps and Continuous Delivery," 2017, retrieved: May 2020. [Online]. Available: https://www.microfocus.com/media/ebook/Software-DevOps-eBook.pdf

[33] J. Humble and D. Farley, Continuous Delivery: Reliable Software Releases through Build, Test, and Deployment Automation (Adobe Reader). Pearson Education, 2010.

[34] M. L. Souca, "Customer dissatisfaction and delight: completely different concepts, or part of a satisfaction continuum?" Management & Marketing, vol. 9, no. 1, 2014, pp. 75–90.

[35] M. Zeelenberg and R. Pieters, "Beyond valence in customer dissatisfaction: A review and new findings on behavioral responses to regret and disappointment in failed services," Journal of business Research, vol. 57, no. 4, 2004, pp. 445–455.

[36] H. Zeitoun and E. Chéron, "Mesure et effets de l'insatisfaction: application au marché des services aériens," Recherche et Applications en Marketing (French Edition), vol. 5, no. 4, 1990, pp. 71–86.

[37] J. Goodman and S. Newman, "Understand customer behavior and complaints," Quality Progress, vol. 36, no. 1, 2003, pp. 51–55.

[38] S. W. Rayburn, "Consumers' captive service experiences: it's you and me," The Service Industries Journal, vol. 35, no. 15-16, 2015, pp. 806–825.

[39] O. Furrer, "La satisfaction des clients des services captifs," vol. 2018-2, 10 2018, pp. 51–75.

[40] TU Dresden, ZIH, "TU Dresden, ZIH, IT services," September 2018, retrieved: May 2020. [Online]. Available: https://tu-dresden.de/zih/dienste

[41] International Telecommunication Union/ITU Telecommunication Sector, "Standard ITU-T E.440: Terms and Definitions Related to Quality of Service and Network Performance Including Dependability – Telephone Network and ISDN Quality of Service, Network Management and Traffic," 1996, retrieved: May 2020. [Online]. Available: https://standards.globalspec.com/std/704295/itu-t-e-440/

International Journal on Advances in Intelligent Systems, vol 13 no 1 & 2, year 2020, http://www.iariajournals.org/intelligent_systems/

Music Event Detection Leveraging Feature Selection based on Ant Colony Optimization

Jian Xi*†, Michael Spranger† and Dirk Labudde†‡
†University of Applied Sciences Mittweida
Forensic Science Investigation Lab (FoSIL), Germany
Email: {xi, spranger}@hs-mittweida.de ‡Fraunhofer
Cyber Security
Darmstadt, Germany
Email: labudde@hs-mittweida.de

Abstract—**Announcements of events are regularly spread using the Internet, e.g., via online newspapers or social media. Often, these events involve playing music publicly that is protected by international copyright laws. Authorities entrusted with the protection of the artists' interests have to find unregistered music events in order to fully exercise their duty. As a requirement, they need to find texts in the Internet that are related to such events like announcements or reports. However, event detection is a challenging task in the field of Text Mining due to the enormous variety of information that needs to be considered and the large amount of data that needs to be processed. In this paper, a process chain for the detection of music events incorporating external knowledge is proposed. Furthermore, a feature selection algorithm based on ant colony optimization to find featurse with a high degree of explanatory power is presented. Finally, the performance of five different machine learning algorithms including two learning ensembles is compared using various feature sets and two different datasets. The best performances reach an F_1-measure of 0.95 for music texts and 0.968 for music event texts, respectively.**

Keywords–Event Detection; Text Classification; Named Entity Recognition; Feature Selection; Ant Colony Optimization.

I. Introduction

In a highly connected world, in order to assure the rights of artists, it is of utmost importance to develop an automatized solution to retrospectively detect violations against copyrights and exploitation rights related to music events. In [1], a first approach towards such a system for online data was proposed. This paper deepens the discussion and proposes an additional feature selection method based on ant colony optimization (ACO).

Individuals, groups and organizations can infringe artists' copyrights in different ways. Whereas individuals might create unauthorized copies, groups and organizations responsible for public events or artists playing at them might play music or show movies without respecting the copyright interests of artists, either deliberately or due to ignorance of the law. Pursuing individual interests of artists is difficult to realize due to practical reasons (e.g., lack of information). Therefore, authorities or private institutions are entrusted with the artists' interests. Usually, organizers register music or movie titles, which are going to be played at an event together with the expected number of participants with these representatives and by paying for the licenses receive the right to play these titles. The official authorities and institutions of the artists then transfer the license fees to the respective artist. One of the largest private institutions in Germany is the Gesellschaft für musikalische Aufführungs- und mechanische Vervielfältigungsrechte (GEMA, English: Society for musical performing and mechanical reproduction rights) representing the rights of about 2 Million artists all over the world and with a total revenue of 1 Billion Euros a year [2]. So far, finding unregistered events after they have taken place is very difficult and is a process mostly done manually.

Nowadays, the information that an event is taking place is often spread using online newspapers, Facebook, Twitter as well as websites. Additionally, after an event has taken place it is often discussed using the same means of communication. Spreading the information this way is often the first choice, as many people can be reached in a short amount of time. Hence, analyzing these textual data makes it possible to automatically find the information needed to uphold the artists' rights. Text Mining, also referred to as Text Analysis, focuses on the analysis of texts in order to receive high level information and latent patterns. It can be applied to many different areas [3], however, it plays a special role in forensics and predictive policing, where it can be used to detect events with a potential to escalate [4]. Event detection is a specific Text Mining problem in which texts are analyzed in order to mine a set of texts that have a semantic link or share conceptual patterns regarding past or future events [5] [6].

The study specifically addresses the detection of music events announced or talked about in social media and online newspapers, with the intention to find those events where copyrights might be violated. Because of the vast amount of data that needs to be taken into account, the data can only be effectively analyzed using machine learning techniques and methods applied in automatized text classification [7]. Since the selection of features is a critical step due to the "curse of dimensionality" problem [8], the Ant Colony Optimization (ACO) algorithm was used for feature selection. To the best of our knowledge, there have been no studies, so far, using ACO for event detection problems. As shown in Section V-B, when using ACO for the selection of features, the experimental results slightly improve compared to the results achieved using the approach proposed in [1]. Additionally, the number of features representing the documents decrease dramatically.

This paper is organized as follows: In Section II, some related work is briefly reviewed. Section III describes difficulties in the current domain as well as the development of a gold standard. The proposed concept is explained in Section IV. Details about the experimental evaluation, the results as well as a short discussion can be found in Section V. Finally, in Section VI a short conclusion is given and some aspects of future work are discussed.

II. Related Work

As mentioned above, event detection is a special text mining problem and can also be seen as a classification problem [9]. However, first it needs to be clarified how an event is defined. In this study, we chose a definition based on the ontology by [10]. Accordingly, an event is defined by the presence of agents at a specific time and place, who are engaged with or in a common matter (product) under concomitant circumstances (factors). Event detection was initiated by the Topic Detection and Tracking research program [11], yet, only focuses three of the original five tasks, namely: tracking, detection and first story detection [12].

In the past years, several approaches have been developed for closed and open domains. For the former manually designed keyword lists can be used to detect specific events in texts [13]. Those keyword lists work effectively, yet need expert knowledge to define the event-specific keywords. Furthermore, keyword lists are limiting the search framework, which is why they will not work for open domains and can only be used as an additional resource for more complex event types, as is the case with the detection of music events. Another example for the detection of events within a specific field is presented by [14] and [15], both working on the detection of economic events that might influence the market, such as mergers. For open domains, [6] proposed a method using machine learning techniques, like clustering and Named Entity Recognition (NER) combined with an ontology (DBpedia) in order to classify Tweets into eight predefined event categories.

Similar to event extraction, the recognition of events might also be categorized as data-driven or knowledge-driven event recognition. In [14] and [15], data-driven approaches were used, both taking mentions of real-world occurrences into account in order to classify their texts into different types of economic events. However, the data-driven approaches fail to consider semantics. In contrast, knowledge-based approaches focus on mining patterns from data to deliver potential rules representing expert knowledge. Depending on the domain or the context, linguistic, lexicographic as well as human knowledge or a combination of these is applied [16].

Much work has been done concerning event detection using different approaches within different fields. Certainly, some of the proposed methods, such as those presented in [13] and [6], can be applied for the detection of music events and our concept is based on the work by [6]. However, in the domain of music event detection, some difficulties appear. For example, events might be announced only using the name of an artist. Some of these difficulties will be discussed in later sections.

One example for a study on music events working with Twitter data is given in [17]. In their study, they identify musical events mentioned in Twitter in order to create a list including sets of artists and venues. The information can be added to an already existing list, for example, a city event calendar [17].

III. Data Preparation

Since the nature of the data is very heterogeneous – different sources like Facebook and newspapers are considered – its analysis has inherent challenges. Below, some of them are discussed in more detail.

A. Data Sources

At the beginning of the study, experts, during their work on manually detecting unregistered music events, independently and arbitrarily preselected more than 1000 music event relevant and irrelevant texts from Facebook and online newspapers. This dataset was then annotated as presented below and used as a basis for our gold standard.

B. Challenges

Noisy Data: In general, texts from social media are inherently characterized by noise. For example, texts often include web addresses, telephone numbers, dates and other characters like hashtags. Furthermore, the texts posted, for example, on Facebook or Twitter are not well written in terms of their grammar and orthography. The application of standard NLP tools to correct such mistakes may lead to incorrectly written names of musicians. As these names are crucial for this study, important events may not be detected.

Text Length: Due to technical restrictions and their intended usage, texts in social media are often very short. Information is compressed as much as possible, for example, by using emoticons or abbreviations or by completely leaving out words. Therefore, the application of standard text analysis methods is often difficult, especially, if the method relies on syntactically correct structures. Considering the following text from Facebook, the application of standard Named Entity Recognition methods fails, because some syntactic features are missing:

"Foo Fighters Eintritt 19. in Hamburg"

Latent Information: Taking the example from above, the crucial information that needs to be found is – even if the text is already classified as an event – that Foo Fighters is a band name and, therefore, the text announces a music event. Typically, such information is extracted by applying methods from the field of NER as discussed in [18]. Traditionally, NER is a subtask in the field of information extraction that focuses on locating structured information in a text and assigning it to predefined categories such as names of persons, organizations and locations. However, distinguishing normal persons from singers or normal organizations from bands is challenging and presents one of the biggest problems in the selection of appropriate features as no prior information is available that indicates whether what the NER model identified is really music-related. This can be changed by adding additional information in the gazetter. This means, before the classification it is already known that, i. e., Johann Sebastian Bach is a musician. However, a much more challenging task is the identification of entities in a text such as musicians that are unknown, for example, a new band or DJ. Unfortunately, texts including these entities appear more often than texts announcing events with known entities.

Dynamic Entities: Information is always dynamic and changes in meaning depending on the time of production. The latent new NER-entities (e.g., musicians, bands or groups) change over the time. An example would be the singer and songwriter Ed Sheeran. Before he became a known musician, he would need to have been labeled as a normal person. However, now he needs to be labeled as a musician. This means, which named entities are relevant changes depending on the point of time a text was written. This triggers the requirement to simultaneously update the knowledge base of our system.

C. Gold Standard

Because there are no suitable training data available, it was necessary to create a gold standard as a basis for the training and evaluation of various classification models. As was mentioned above, texts were collected arbitrarily, including 21 texts from online newspapers and 1,097 texts from Facebook. These were manually annotated as music related or music unrelated as well as event related or event unrelated. Both decisions were made independently of each other. Due to text-inherent vagueness, the data was independently labeled by 35 people. In order to ensure the quality of the labeled data, each person was only allowed to work for 2 hours a day.

The final decision regarding what category a text belongs to was made by using a majority criterion. This criterion requires a minimum number of people to agree on a decision in order to provide a confident classification. If the minimum number of agreements was not achieved for a given text, the text was considered ambiguous and removed from the corpus. The minimum number of agreements was derived from a binomial test under the null hypothesis that each decision individually made by every study participant is conducted at random. This hypothesis thus states that $p^+ = p^- = 0.5$, where p^+ and p^- are the decision probabilities. With respect to the null hypothesis, for every number of agreements d a probability $P(d|p^+)$ can be derived from the corresponding binomial distribution. The minimum number of agreements d_{crit} is equal to d, where the null hypothesis can be rejected according to $P(d \geq d_{crit}|p^+) < \alpha$. Here, α corresponds to the Bonferroni-corrected significance level of $0.05/n$, with n being the number of considered texts. In this study, the minimum number of agreements d_{crit} was 29 for the text corpus.

As a result, the corpus consists of 19 newspaper texts and 867 Facebook texts. 335 out of the 867 Facebook texts and 14 out of the 19 newspaper texts are music relevant. Table I provides some descriptive statistics. When music event classification is considered, the number of texts that meet the Bonferroni constraint drops to 505, whereas 251 Facebook texts and 9 online newspaper texts are music event relevant. Table II provides the descriptive statistics for the music event related data. In summary, at the end, two datasets were created: one for music relevance, including 886 texts and one for music event relevance with 505 texts.

TABLE I. STATISTICS OF THE DATA REGARDING MUSIC DETECTION.

	# texts	$\#_{tot}$ words	$\#_{avg}$ words	shortest	longest
newspaper	19	2,071	109	14	387
Facebook	867	85,965	99.1	1	1,238
total	886	87,965	99.3	1	1,238

TABLE II. STATISTICS OF THE DATA REGARDING MUSIC EVENT DETECTION.

	# texts	$\#_{tot}$ words	$\#_{avg}$ words	shortest	longest
newspaper	13	1,077	82.85	14	277
Facebook	492	59,440	120.81	1	1,238
total	505	60,517	119.84	1	1,238

In order to describe the data in the domain of music events, we defined an XML-schema, with which our raw data can be concisely structured in order to serve as a gold standard to train and test models in this field. Even though this work is focused on music event detection, the schema is constructed to contain various types of event data, such as music, theater, or readings. It includes, beside others, the following information:

- raw text
- source (e. g., Facebook)
- event-related ($\{0, 1\}$ and certainty)
- event-type-related ($\{0, 1\}$ and certainty)
- event location
- event-date
- persons
- different types of roles (e. g., musician, actor)
- different types of events (e. g., music, theater)

It needs to be emphasized that the relation between any text and a specific category is described twice: binary and with a numeric value. The binary description refers to the classification and thus serves as a ground truth, whereas the numeric value represents the degree of certainty. With this gold standard the following areas may be addressed:

- classification of texts regarding different event-types
- recognition of event-related entities, i. e., roles of persons, organizations and locations

Named entities are considered because they provide strong features for the classification, as was shown in [19] and [20]. For example, if the name Eric Clapton, an English singer and songwriter, appears in a text, this is a strong indication that the current text is music related. Since classic NER mostly concentrates on distinguishing between persons, locations, and organizations, a more detailed categorization including some kind of prior knowledge is needed. The entire dataset was annotated and curated manually according to the schema described so far.

IV. Proposed Concept

The task of detecting texts concerning music events is a typical categorization task. Categorization, as a special case of classification, attempts to categorize a text into a predefined set of conceptual categories using machine learning techniques. Formally, let $T = t_1, ..., t_m$ be a set of texts to be categorized, and $C = c_1, ..., c_n$ a set of categories, then the task of categorization can be described as surjective mapping $f : T \to C$, where $f(t) = c \in C$ yields the correct category for $t \in T$. In the field of music event detection, texts need to be assigned to one out of two main classes: related to a music event or not. Texts of the former class can be further categorized into different event types, such as public concerts. This might be of great importance as some music, e. g., religious music or classical music concerts, are license free or public music resources.

Currently, institutions responsible for the enforcement of exploitation rights have to detect unannounced music events predominantly manually and with the help of search engines. This leads to various problems. Firstly, the manual search is very inefficient on large-scale data. Secondly, the manual checking process is error-prone and differs depending on the person who judges the data. Furthermore, the current process chain can hardly be deployed in an online mode due to its semi-automated nature.

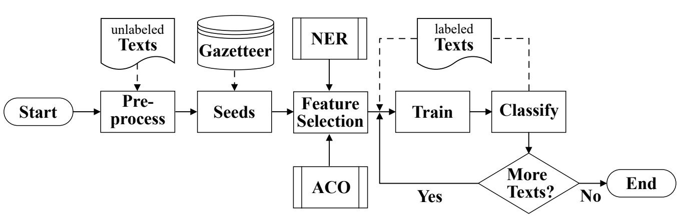


Figure 1. The proposed workflow of music event detection.

To overcome these limitations, a semi-supervised process-chain using a bootstrapping approach, as depicted in Figure 1, is proposed. The advantage of the chosen approach is that the training can start with very few but highly descriptive examples in order to create a first restrictive classifier which will be further improved in upcoming iterations until all texts are classified or no further improvement is possible. Next, each step is discussed in more detail.

A. Preprocessing

As mentioned in Section III-B, the texts we worked on mostly come from the Internet. Such texts often contain typing errors and are often written in informal language, including dialect. This leads to even noisier data than usual in textual texts. Besides common shallow text preprocessing, including stopword and punctuation removal as well as stemming or lemmatizing, there is a strong need for additional language information. This information can be provided in the form of a knowledge base curated by experts. For instance, preselected terms, such as party or live music, can be used to build the gazetteer. Additional useful information might be venues of interest, such as clubs or cafés, where music events often take place. In short, information directly related to music events can be used as a basis of knowledge. This knowledge base can be a simple gazetteer, as is the case in our study, or can incorporate more complex structures, as in [21].

B. Collecting Seed Texts

The most crucial task in bootstrapping is finding seed texts which represent the concept of the classes as well as possible. The usage of some kind of highly descriptive key words or phrases collected from experts in this field is one possible way to find seed texts in a highly accurate, but, nevertheless, very restrictive way.

C. Named Entity Recognition (NER)

As was shown in [18]–[20], named entities might be a useful feature for text classification tasks. In a first step, named entities are identified using any NER method, as discussed in [18]. However, as was already discussed in Section III-B, the named entities detected this way are not specific enough. Hence, domain-specific knowledge resources like MusicBrainz, an open music encyclopedia, and DBpedia can serve as a music database for distinguishing recognized entities further, in order to assign appropriate roles to them, for example, `musician` to a person. The richer and up-to-date this knowledge base, the more accurate is the classification. The entire process of music event related Named Entity Recognition is shown in Figure 2. The influence of using NER with a knowledge base is clearly shown in Section V-B.

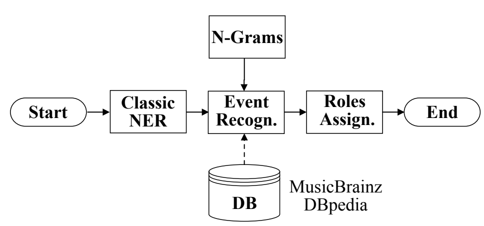


Figure 2. The proposed workflow of detecting music related named entities.

D. Feature Selection

The next step is the selection of appropriate features to represent the text data. Feature selection is always a critical step in text classification tasks. On the one hand, well selected features are necessary to achieve highly accurate results. On the other hand, they help reduce the feature space and, as a consequence, minimize the time complexity [22].

Generally, feature selection, as a typical machine learning task, can be distinguished in supervised, unsupervised and semi-supervised approaches depending on whether the data is labeled or not [23]. The approach for music event detection used in this paper relies, at least initially, on the availability of labeled training data. Thus, it belongs to the algorithmic class of supervised methods. Below, three general approaches in supervised feature selection are briefly introduced:

1) **Filter Approach**: Here, the explanatory power of features is measured by using intrinsic properties of data to select the best feature without employing predictive models. The principal components of methods based on this approach are the feature search and selection criterion. The feature can be ranked according to the score evaluated by statistical measures such as the chi-squared test or the information gain [22]. The ranked features can be either removed or selected from feature sets by comparing their explanatory power with a given threshold. Thus, the most challenging part of this method is to select the proper feature candidates [24].
2) **Wrapper Approach**: Methods based on this approach utilize a machine learning model to evaluate the explanatory power of the elements in the feature set. Hence, the selection of proper features can be considered as a search algorithm. In this case, candidate features are found using a search strategy like heuristic search, sequential forward selection, or backward elimination [24]. The selected features are then used to train a predictive model and evaluate the fitness of the selected features by utilizing hold-out data. The overall procedure is repeated until the desired quality criterion is met. In this approach, the trained model can be seen as a black box and its representational bias is decreased by repeatedly applying the feature search process (i.e., by means of cross validation). The feature selection method based on ACO as proposed in this paper is inspired by this approach.
3) **Embedded Approach**: Here, a linear classifier such as SVM or logistic regression is chosen in order to learn a predictive model and, at the same time, select the most appropriate features. An additional regularization function is usually included in order to constrain the learning of the coefficients of this model. The features, whose coefficient is non-zero, are selected [23].

Ant Colony Optimization Algorithm: Heuristic algorithms and especially those using swarm intelligence, such as ACO, are developed to solve combinatorial optimization problems. They are based on the natural principles of self-organization as a mechanism to control the behavior of the individuals of the swarm. In this sense, the ACO is inspired by observing the behavior of ants in an ant colony and was first reported in [25]. An ant population is able to find a solution for a combinatorial problem, which of course might be sub-optimal, by mimicking the ants' forage by performing randomized walks between food resources and their colony.

This highly coordinated behavior between the individuals in an ant colony can be transferred to solve computational tasks. In this case, each ant iteratively tries to find a possible solution to solve the problem taking into account its own current heuristic information as well as additional information propagated by other ants. In this way, the population as a whole is able to find one, though maybe sub-optimal, consensus solution. Whenever an ant finds a candidate solution, the information about the path this ant has walked is spread throughout the population guiding following individuals. This kind self-organization is made possible by the chemical odoriferous redolence, called *pheromone*, which is left by each individual of the colony.

During the search, each ant contributes proportionally to the final optimal solution. The repeated communication between individuals happens in-directional by changing their environment. Together the pheromone paths, left behind by the interaction between the individuals, form a pattern which is bigger than the pattern of each individual [26]. This collaborative behavior pattern is called *stigmergy* [27].

ACO based Feature Selection: In [28], ACO based feature selection was used to select feature subsets for different disease data sets. An artificial neural network was trained as a predictive model in order to evaluate each feature subset. It is also shown that the selection of feature can improve the performance of a classifier. In a different study a variation of ACO, *AntMiner+*, was used to mine rules that provide information about the decision making process, i.e., considering a couple of given attributes such as amount of bank deposits, duration of the deposits, or credit histories [29]. Consequently, a rule-based relation is extracted in order to classify the customer's creditworthiness [29]. Although the time complexity of ACO is the highest in this study, it still provides plausible results [29].

In the field of text classification, the feature space is usually large. ACO-based feature selection works under the assumption that inter-variable relations among a reduced feature subset represents the original data in a way that the predicted results are accurate [28]. As shown in [30], the ACO-based feature selection outperforms genetic algorithms, information gain, and the chi-squared test on the Reuters-21578 data set. Similar results were achieved by [31] for the classification of web pages using a decision tree, Naïve Bayes and k-NN each combined with an ACO-based feature selection method. Using the same data set as [28], [32] also effectively applied a ACO-based feature selection method. The main difference between the two studies are the strategies for state transitions and the pheromone updates. Subsequently, the feature selection approach used in this study is based on these two studies.

When applying ACO, in a first step the problem should be represented as a graph, whereas each node in the graph represents a single feature in the feature space [25] [28] [29] [32]. In this study, all the tokens are stemmed by using the same pre-processing method as introduced in Section IV-A. Following the principle of ACO, those features selected by the ants, which lead to better test results, get a higher pheromone concentration and higher heuristic values and are, consequently, selected more frequently. Hence, there is no additional weighting for features necessary, however, a strategy for initializing the features in feature selection is needed. Here, we experimented with two strategies, one using a constant value and the other Mutual Information (MI). The features are selected and evaluated according to their desirabilities and contributions to the fitness. The details are shown in Section IV-D. This process is repeated until the desired explanatory power of the resulting feature set is reached.

Algorithm 1 Proposed Feature Selection Algorithm with ACO

Require: All features
Ensure: The best feature subsets S^k_{best}
1: Initial colony information τ, η
2: Generate Ants $\mathcal{A}$
3: $i \leftarrow 0$
4: **while** $i \neq K$-Fold **do**
5:   Prepare training/Test data in i-Fold
6:   **for** $l = 0$ to L-iteration **do**
7:     **for** ant $k \in \mathcal{A}, t \in \mathcal{T}$ **do**
8:       Construct feature subsets $S^k(t)$ in i-Fold
9:     **end for**
10:     **if** all constructions are finished **then**
11:       Evaluate feature subsets $S^k(t)$ in i-Fold
12:       **if** stop condition satisfied **then**
13:         Return the feature subsets S^k_{best}
14:       **else**
15:         Update colony information τ, η
16:         Reset ant memory
17:         Generate Ants $\mathcal{A}$
18:         Go to step 5
19:       **end if**
20:     **else**
21:       Go to step 5
22:     **end if**
23:   **end for**
24:   $i \leftarrow i + 1$
25: **end while**

Algorithmic Details: The feature selection algorithm used in this study is shown in Algorithm 1 and explained in more detail below:

Initializing the Colony: The pheromone level τ and the heuristic information ρ should either be initialized with a constant value as suggested by [28] and [32] or with more informative values like the information gain or the Pearson Correlation Coefficient as suggested by [33]. In this study, the initialization of the pheromone level and the heuristic information is done once assigning a constant value and once assigning MI.

Generating Ants: As suggested in [32], the number of ants $\#ants$ should equal the number of features in the feature space $\#features$. Unfortunately, in our case, there are at least 10,000 features. Therefore, in order to reduce the computational complexity only $\#ants$ = $\#features/100$ were created. Note that we only decreased the number of ants in the ant colony, however, the features to be crawled stayed unchanged.

Constructing the Feature Sets: Each ant has a limited capacity to hold the features it crawled. In [32], the roulette-wheel schema was used to decide the size of the feature subset $\mathcal{T}$ for each ant given a hyper-parameter μ. In this study, μ is set to $\mu = 0.35$. This factor influences how many features an ant can take during its trail constructing process. The next problem is how the ants pick a particular feature (i.e., a node in feature graph) in such a way that the classification accuracy is maximized.

The artificial ants move between features under a *pseudo-random proportional rule*. In other words, a probability decision policy guides the ants' walk through the adjacent features in the search space. In [32], the state transition does not consider exploration but only exploitation, while [28] uses a random number drawn from a uniform distribution in order to control exploration and exploitation. In order to get informative, yet still compact features without losing generalization the same state transition policy as the one in [28] was used in this study.

The probability of a feature i at time step t to be selected by ant k is defined in Equation (1):

$$P_i^k(t) = \begin{cases} \arg\max[\tau_i(t)]^\alpha \times [\eta_i(t)]^\beta, & \text{if } (q < q_0) \\ \frac{[\tau_i(t)]^\alpha \times [\eta_i(t)]^\beta}{\sum_{u \in j^k} [\tau_u(t)]^\alpha \times [\eta_u(t)]^\beta}, & \text{if } (u \in j^k) \end{cases}, \quad (1)$$

where factor α, η controls the influence of the pheromone and heuristic of feature i at time step t on the transition, respectively. The random number $q \in [0, 1]$ drawn from a uniform distribution, controls the trade-off between exploration and exploitation. The hyper-parameter q_0 forms the threshold value for this trade-off and j^k is the set of possible features that can be taken by ant k at time step t.

Checking whether all constructions are finished: the construction step repeats until all ants have finished constructing their tours.

Evaluate Feature Subsets: In this step, each feature subset $S^k(t)$ is used to train an arbitrary machine learning model. Here, the Multi-Layer-Perceptron algorithm is used with the same parameter settings as proposed in former experiments described in [1], where it outperformed all other considered models.

Afterwards, the features are evaluated with hold-out data. The pheromone concentration is proportionally updated depending on the fitness of the features $f(S_k(t))$. Additionally, the fitness of the features also gives information about how good a feature is and, thus, its desirability. In this study, a 3-fold cross validation was used to evaluate the features. Within each fold, out of all the best local feature subsets $S^l(t)$, calculated during L iterations, here $L = 10$, the best global feature subset was determined. During an iteration, if the fitness of the selected best local features does not change anymore, the iteration breaks.

This way, the colony does not only focuses its search for optimal features in a fraction of the dataset, yet in the whole training dataset, i.e., feature space [32]. The best local and global feature subset contribute to the fitness of features as is explained in the following step.

When a pre-defined stop criterion is reached, the best feature subset will be returned and the procedure is terminated. Otherwise, the colony information such as pheromone and heuristic of features is updated with respect to their contributions.

Updating the Colony Information: Based on the fitness of features determined during the previous feature evaluation step, the pheromone values and the heuristic information of each feature are updated. In this study, the same update rules as in [32] were applied (see Equation (2)).

$$\tau_i(t+1) = (1-\rho) \times \tau_i(t) + \frac{1}{m_i} \sum_{k=1}^{\#_{ant}} \Delta_i^k(t) + e \times \Delta_i^g(t), \quad (2)$$

where ρ is the evaporation parameter of the pheromones, m_i the absolute frequency of the feature i considering all feature subsets and e the elitist parameter which decays the contribution of features occurring in the best local feature subsets. $\Delta_i^k(t)$ is defined as the fitness of those features that are selected by the ants during their current tour and is calculated as shown in Equation (3).

$$\Delta_i^k(t) = \begin{cases} f(S^k(t)) & \text{if } i \in S^k(t) \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

$\Delta_i^g(t)$ is defined as the sum of the fitness of those features that occur in the best local feature subsets. Equation (4) shows how it is calculated.

$$\Delta_i^g(t) = \begin{cases} f(S^l(t)) & \text{if } i \in S^l(t) \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

Subsequently, both results are normalized by their occurrence frequency. The decayed pheromone of each feature is incrementally updated by considering local updates from all tours and the penalized global updates from the elitist tours. The heuristic of each feature is updated as shown in Equations (5) and (6).

$$\eta_i(t+1) = (1-\rho) * \eta_i(t) + \Delta_{\eta_i}, \quad (5)$$

where

$$\Delta_{\eta_i} = \frac{1}{m_i} \sum_{k=1}^{\#_{ant}} f(S_i^k(t)) \times \left[1 + \phi \times \exp\left(-\frac{|S_i^k(t)|}{\#_{feature}}\right)\right], \quad (6)$$

and ϕ weights the exponential ratio between $|S_i^k(t)|$, the length of selected feature subset by k in step t in which the feature i occurs, and the total size of the feature space. Those features that are associated with a reduced subset and yield a better test performance than other features deserve a higher desirability.

Reset Ant Memory: All the features selected by ants, their pheromone concentrations as well as their desirability are released for the next tour.

This procedure can be finished with $\mathcal{O}(ikn)$ time complexity and $\mathcal{O}(kn^2)$ space complexity, where i is the iteration, k the number of ants in the colony and n the number of features. It can be noticed that, in this study the ACO approach includes, in addition to the pheromone level τ and the heuristic information η other hyper-parameters, namely, $\mu, \alpha, \beta, \rho, e, \phi, q_0$. While [27] discusses how the pheromone level and the heuristic information can be determined, so far, there have been no studies reporting how the additional parameters are set. Therefore, the optimal settings for these parameters have to be determined empirically.

E. Training and Classification

The final step is to train a first classifier using the seed texts and to try to assign categories to the other texts. This step is repeated until no improvement of the classifier can be achieved or no remaining texts are left.

F. System Complexity

In the following section, the system complexity shall be briefly described on the basis of time and space.

Time Complexity: The time complexity of the system, without considering the training and classification process, can be described as shown in Equation (7),

$$\begin{aligned} T(n) &= T_{pre} + T_{gazetteer} + T_{seeds} + T_{ner} + T_{fea_{sel}} \\ &= \mathcal{O}(2^p) + 2\mathcal{O}(1) + 3\Theta(lp) + \mathcal{O}(L|S|^3) + \mathcal{O}(ikn) \end{aligned} \quad (7)$$

where L is the number of samples and $|S|$ the number of labels in the NER process. Furthermore, l is the length of the string, p the length of the search pattern in the string, i the iteration in feature selection, k the number of ants in the ant colony, and n the number of features.

Space Complexity: Similarly, the space complexity can be measured without considering the training and classification process as shown in Equation (8),

$$\begin{aligned} T(n) &= T_{pre} + T_{gazetteer} + T_{seeds} + T_{ner} + T_{fea_{sel}} \\ &= \mathcal{O}(2^p) + \mathcal{O}(g) + \Theta(lp) + \mathcal{O}(s+l) + 2\Theta(lp) \\ &\quad + \mathcal{O}(r) + \mathcal{O}(kn^2) \end{aligned} \quad (8)$$

where g is the size of the gazetteer, r the number of roles, s the size of the trained NER-model, k the number of ants in the ant colony and n the number of features. After analyzing the time and space complexity, it can be shown that the system requires intensive resources in preprocessing and in identifying named entities with respect to time and space complexity. Thus, the performance of our system, regarding time and space complexity, depends on the methods that are used in these two setups.

V. Experimental Evaluation

To create first baseline results, the labeled data (see Section III-C) were categorized using the following supervised machine learning methods: Naïve Bayes, SVM, and MLP and two ensemble approaches: AdaBoost and RandomForest [34]. The categorization was done once with each dataset. Firstly, the dataset with 886 texts was used and categorized as music relevant or not. However, as the ultimate goal is a system for the detection of music events and not just music, secondly, the dataset with only 505 texts was categorized as music event relevant or not.

The performance of each experimental setting was determined using micro-averaged precision (Micro P.) and recall (Micro R.) as well as the F1 measure.

A. Setup

In this study, only two sources of texts concerning music events are considered: Facebook as well as daily and weekly online newspapers. The raw data were preprocessed as described in Section IV-A. Furthermore, all numbers, for example, telephone numbers and dates, were removed and, therefore, not considered in the categorization. For comparison, two different datasets for each dataset were created. The first dataset contains word tokens that were processed with the Porter stemmer [35], whereas for the second dataset the algorithm proposed in [36] was used. For the detection of named entities a Conditional Random Field approach was applied, as proposed by [37]. As was mentioned in Section IV-D, MusicBrainz und DBpedia were used to assign roles to named entities and were combined in order to increase the number of matches.

In this study, the following four representations of the texts incorporating different features were compared:

- multinomial Bag of Words (BoW),
- Term Frequeny Inverse Document Frequency representation of BoW (TF-IDF(BoW)),
- multinomial BoW and music event related Named Entities (BoW+NE), and
- TF-IDF representation of BoW and NE (TF-IDF(BoW+NE)).

In case of named entities only their type (role) was considered as a feature rather than the entity itself, e. g., song writer or musician were taken as a feature instead of Eric Clapton. Moreover, it was only possible to train the SVM with frequency-based features. In the ensemble approaches, an SVM (using the same setting as reported in [1]) and a decision tree were used as basic classifiers. This is distinguished later by *svm* and *tree*, respectively. As a criterion to measure the quality of split, "Gini" for the Gini impurity and "Entropy" for the information gain were considered, which are labeled with *gini* and *ent*, respectively.

B. Results

Because the features were investigated by means of ACO by using a 3-fold cross validation due to the computation time, the same experimental settings as in [1] were used. Additionally, for the ACO-based selection of features the following settings for the hyper-parameter were used: $\alpha = 0.5$, $\beta = 0.8$, $\rho = 0.5$, $q_0 = 0.6$, $e = 0.6$, $\phi = 0.5$, $\omega = 0.5$. The original number of features for both datasets used in this study is shown in Table III, while the size of the selected features is shown in Table IV. For initializing the pheromone and heuristic in ACO, a constant value for the first experiments was used and later MI.

TABLE III. NUMBER OF FEATURES OF THE MUSIC AND THE MUSIC EVENT DATASET.

	# raw	# raw including NEs
Music Relevance	14,275	10,921
Music-Event Relevance	12,171	9,268

TABLE IV. NUMBER OF SELECTED FEATURES OF THE MUSIC AND THE MUSIC EVENT DATASET.

Dataset	Initialization	# raw	# raw including NEs
Music Relevance	Constant Value	11,164	9,584
	MI	9,921	8,428
Music-Event Relevance	Constant Value	8,618	7,375
	MI	7,852	6,485

As can be seen in Table IV, the number of features is considerably reduced in both datasets after using ACO to select the features. The overall results achieved with this reduced feature set when being used in the proposed classification pipeline is shown below:

Baseline Results: The baseline results of the music relevance decisions regarding the gold standard dataset described in Section III-C are shown in Table V whereas the results for the categorization of music event relevance are shown in Table VI. The results using stemming were compared with those achieved using lemmatization and it was observed that stemming lead to slightly better results. As can be seen in Table V, the best results for the categorization of music relevance, based on the F_1-measure, were achieved using a frequency-based representation of words and named entities (roles) and MLP. Comparable results were achieved by an AdaBoost model based on SVMs as the basic classifier but considering the same feature settings. In comparison, a combination of BoW and named entities (roles) and an MLP model achieved the best results for the categorization of music event relevance. These results are presented in Table VI. Furthermore, it was found that the best performing model and feature combination (MLP and BoW+NE) failed when the features (words) occurred in both, the relevant and non-relevant texts, when the texts were very short or when there were not enough features with a high explanatory power available.

Overall, the classification results of music relevance and music event relevance are clearly improved when named entities are considered as features.

Considering ACO Feature Selection: As mentioned in Section IV-D, by using MI to initialize the pheromone and heuristic of the features, 9921 features were found using the 886 raw documents and 8428 features using 886 documents by considering named entities in the documents. This is repeated

TABLE V. RESULTS OF THE 3-FOLD CROSS VALIDATION FOR THE MUSIC DATASET USING STEMMING.

Model	Feature	Micro P.	Micro R.	F_1
Naïve Bayes	BoW	0.656	0.974	0.784
	TF-IDF(BoW)	0.991	0.625	0.766
	BoW+NE	0.699	**0.983**	0.817
	TF-IDF(BoW+NE)	0.988	0.736	0.844
MLP	BoW	0.886	0.874	0.880
	TF-IDF(BoW)	0.884	0.897	0.890
	Bow+NE	0.943	0.903	0.922
	TF-IDF(BoW+NE)	0.934	0.931	**0.933**
SVM	TF-IDF(BoW)	0.961	0.840	0.896
	TF-IDF(BoW+NE)	0.975	0.894	0.933
AdaBoost	*svm*_TF-IDF(BoW)	0.961	0.840	0.896
	*svm*_TF-IDF(BoW+NE)	0.975	0.894	**0.933**
	*tree*_TF-IDF(BoW)	0.982	0.633	0.770
	*svm*_TF-IDF(BoW+NE)	**0.996**	0.719	0.835
RandomForest	*gini*_TF-IDF(BoW)	0.985	0.771	0.865
	*gini*_TF-IDF(BoW+NE)	0.990	0.851	0.915
	*ent*_TF-IDF(BoW)	0.981	0.751	0.851
	*ent*_TF-IDF(BoW+NE)	0.990	0.831	0.903

TABLE VI. RESULTS OF THE 3-FOLD CROSS VALIDATION FOR THE MUSIC EVENT DATASET USING STEMMING.

Model	Feature	Micro P.	Micro R.	F_1
Naïve Bayes	BoW	0.892	0.943	0.917
	TF-IDF(BoW)	0.852	0.943	0.895
	BoW+NE	0.907	**0.966**	0.936
	TF-IDF(BoW+NE)	0.873	**0.966**	0.917
MLP	BoW	0.922	0.901	0.912
	TF-IDF(BoW)	0.905	0.909	0.907
	Bow+NE	0.953	0.924	**0.938**
	TF-IDF(BoW+NE)	0.932	0.932	0.932
SVM	TF-IDF(BoW)	0.927	0.920	0.924
	TF-IDF(BoW+NE)	0.946	0.928	0.937
AdaBoost	*svm*_TF-IDF(BoW)	0.927	0.920	0.924
	*svm*_TF-IDF(BoW+NE)	0.946	0.928	0.937
	*tree*_TF-IDF(BoW)	0.965	0.734	0.834
	*svm*_TF-IDF(BoW+NE)	**0.991**	0.795	0.882
RandomForest	*gini*_TF-IDF(BoW)	0.956	0.821	0.883
	*gini*_TF-IDF(BoW+NE)	0.979	0.882	0.928
	*ent*_TF-IDF(BoW)	0.948	0.829	0.884
	*ent*_TF-IDF(BoW+NE)	0.978	0.859	0.915

by constantly initializing the pheromone and heuristic of the features. Table V shows that with dramatically reduced features, the best F_1-measure for the music relevance dataset is improved from 0.933 (in i.e., MLP and TF-IDF (BoW+NE)) to 0.950 (i.e., SVM and TF-IDF (BoW+NE)) when initializing the pheromone and heuristic with MI. Furthermore, all the results are clearly improved by using MI in the initialization, as shown in Table VII. In comparison, the best results do not change dramatically when initializing ACO in a constant way, as can be seen in Table VIII. However, the feature size is clearly reduced in comparison to the original feature size in Table III.

The results in Tables VII and IX show that with the ACO feature selection the performance of all machine learning algorithms are improved: using MI for initializing the pheromone and heuristic, the best baseline result for the music relevance decision is improved from 0.933 to 0.950, and for the music event relevance decision from 0.938 to 0.968.

By analyzing the results of detecting music-event relevance by initializing colony information with MI, the following conclusions were drawn:

a) The classification result can be clearly improved by reducing the feature dimension: the average number of tokens that occur in the documents that are correctly classified after feature selection changes from 73 to 46 and the average classification accuracy is improved from 0.938 to 0.954.
b) The selected features are informative for the classification as shown by the results in Table IX.
c) The false negative classified documents are still classified incorrectly after the selection of features.

It seems that the last mentioned conclusion depends on the distribution of the features in the documents. It was noticed that some music related features lead to unexpected results during the classification. For example, considering the following document:

> "Marienmünster am Freitag. Abtei Marienmünster, 15 Uhr weihnachtliche Orgelmusik, Arien und Instrumentalstücke zum Fest.",

words like *Orgelmusik* (organ music) and *Instrumentalstücke* (instrumentals) give a strong feeling that this document is somehow related to a music event in public. However, as the features only occur once in the corpus, they fail to be considered as strong features by the model to make a decision.

TABLE VII. RESULTS OF THE 3-FOLD CROSS VALIDATION FOR THE MUSIC DATASET USING STEMMING AND ACO FEATURE SELECTION WITH MI INITIALIZED PHEROMONE AND HEURISTIC.

Model	Feature	Micro P.	Micro R.	F_1
Naïve Bayes	BoW	0.811	0.983	0.889
	TF-IDF(BoW)	**0.990**	0.817	0.895
	BoW+NE	0.830	**0.994**	0.905
	TF-IDF(BoW+NE)	0.981	0.885	0.931
MLP	BoW	0.948	0.894	0.920
	TF-IDF(BoW)	0.915	0.926	0.920
	BoW+NE	0.972	0.911	0.941
	TF-IDF(BoW+NE)	0.935	0.951	0.943
SVM	TF-IDF(BoW)	0.969	0.897	0.932
	TF-IDF(BoW+NE)	0.967	0.934	**0.950**
AdaBoost	*svm*_TF-IDF(BoW)	0.969	0.897	0.932
	*svm*_TF-IDF(BoW+NE)	0.967	0.934	**0.950**
	*tree*_TF-IDF(BoW)	0.987	0.645	0.780
	*svm*_TF-IDF(BoW+NE)	**0.996**	0.722	0.837
RandomForest	*gini*_TF-IDF(BoW)	0.993	0.777	0.871
	*gini*_TF-IDF(BoW+NE)	0.991	0.860	0.923
	*ent*_TF-IDF(BoW)	0.983	0.762	0.859
	*ent*_TF-IDF(BoW+NE)	0.990	0.862	0.922

VI. CONCLUSION AND FUTURE WORK

In this paper, two gold standard datasets for music event detection were presented and made publicly available here

TABLE VIII. RESULTS OF THE 3-FOLD CROSS VALIDATION FOR THE MUSIC DATASET USING STEMMING AND ACO FEATURE SELECTION, INITIALIZED WITH A CONSTANT VALUE FOR PHEROMONE AND HEURISTIC.

Model	Feature	Micro P.	Micro R.	F_1
Naïve Bayes	BoW	0.667	0.974	0.792
	TF-IDF(BoW)	0.979	0.673	0.798
	BoW+NE	0.684	**0.986**	0.808
	TF-IDF(Bow+NE)	**0.989**	0.765	0.863
MLP	BoW	0.915	0.862	0.888
	TF-IDF(BoW)	0.904	0.891	0.899
	Bow+NE	0.946	0.903	0.924
	TF-IDF(Bow+NE)	0.921	0.940	**0.930**
SVM	TF-IDF(BoW)	0.964	0.840	0.897
	TF-IDF(Bow+NE)	0.981	0.885	0.931

TABLE IX. RESULTS OF THE 3-FOLD CROSS VALIDATION FOR THE MUSIC EVENT DATASET USING STEMMING AND ACO FEATURE SELECTION WITH MI INITIALIZED PHEROMONE AND HEURISTIC.

Model	Feature	Micro P.	Micro R.	F_1
Naïve Bayes	BoW	0.945	0.985	0.965
	TF-IDF(BoW)	0.938	0.981	0.959
	BoW+NE	0.949	0.985	0.966
	TF-IDF(Bow+NE)	0.932	**0.989**	0.959
MLP	BoW	0.956	0.909	0.932
	TF-IDF(BoW)	0.957	0.939	0.948
	Bow+NE	0.965	0.943	0.954
	TF-IDF(Bow+NE)	0.954	0.951	0.952
SVM	TF-IDF(BoW)	0.966	0.970	**0.968**
	TF-IDF(Bow+NE)	0.966	0.958	0.962
AdaBoost	*svm*_TF-IDF(BoW)	0.966	0.970	**0.968**
	*svm*_TF-IDF(Bow+NE)	0.966	0.958	0.962
	*tree*_TF-IDF(BoW)	0.973	0.696	0.812
	*svm*_TF-IDF(Bow+NE)	**0.986**	0.795	0.880
RandomForest	*gini*_TF-IDF(BoW)	0.965	0.844	0.901
	*gini*_TF-IDF(Bow+NE)	0.975	0.882	0.926
	*ent*_TF-IDF(BoW)	0.978	0.837	0.902
	*ent*_TF-IDF(Bow+NE)	0.975	0.894	0.933

TABLE X. RESULTS OF THE 3-FOLD CROSS VALIDATION FOR THE MUSIC EVENT DATASET USING STEMMING AND ACO FEATURE SELECTION, INITIALIZED WITH A CONSTANT VALUE FOR PHEROMONE AND HEURISTIC.

Model	Feature	Micro P.	Micro R.	F_1
Naïve Bayes	BoW	0.882	0.935	0.908
	TF-IDF(BoW)	0.865	0.947	0.904
	BoW+NE	0.904	0.966	0.934
	TF-IDF(Bow+NE)	0.874	**0.973**	0.921
MLP	BoW	0.912	0.905	0.908
	TF-IDF(BoW)	0.882	0.935	0.908
	Bow+NE	0.961	0.932	0.946
	TF-IDF(Bow+NE)	0.940	0.947	0.943
SVM	TF-IDF(BoW)	0.912	0.901	0.906
	TF-IDF(Bow+NE)	**0.976**	0.935	**0.955**

[38]. Furthermore, a process chain for the categorization of music event related texts was proposed and a first baseline evaluation conducted. For finding much more representative features for event detection, a couple of seed words are given for learning extended features correspondingly. The results change slightly in comparison to results by using the same working process. Further experiments with ACO-based feature selection show that a frequency-based approach considering music specific named entities performs best together with an SVM model for the classification of music relevant texts and an SVM-based ensemble model for the classification of music event relevant texts. As a strategy for initializing the relevant parameters for ACO (e.g., heuristic and pheromone), using MI gives promising results. As discussed in Section V-B, the documents, in which the occurrence of non-domain related features is more dominant than the one of domain related features, are intended to be classified as non-music and event relevant, although these documents are truly music and event relevant. For this purpose, the weights of the features should be considered in classification.

Obviously, the proposed event detection approach can also be applied to similar domains such as movie showings, however, it should be analyzed how it may be applicable to other more general types of events such as social events spread via a social network, or economic events in market and event copyright violation issues. In future work, the following aspects need to be given more attention:

Gold standard datasets: The datasets used in this study are relatively small, especially the one including music event related texts and need to be extended in the future. Alternatively, classification results may be improved without extended datasets by considering further strategies. For example, transfer learning enables the model to use pre-trained knowledge to transfer it to the original problem domain, where there is no sufficient training data available [39]. Based on neural probabilistic language models [40], the texts can be represented by using pre-trained vectors of words that enable the models to observe the semantic in the sentences. Furthermore, active learning shows the ability to reduce the training samples as reported in [41].

Named Entities as Features: In [20], the authors used named entities to represent the documents in a boolean model and a vector space model. Similar work was conducted in [19], where the entity power coefficient is used to create occurrences of all terms that related to a named entity. For the study at hand it would be important to distinguish the different named entities into hierarchical classes, since some music events are related to music that is not protected by any copyright laws. In such cases, there is a need to distinguish these events from those that might include copyright infringements. For this purpose, a more fine-grained categorization to separate different types of events can be realized by applying hierarchical classification methods, such as discussed in [20] [42].

Hyper-parameter of ACO: As discussed in Section IV-D, the ACO approach is accompanied with hyper-parameters and different strategies for initializing the pheromone and heuristic values. In this study, it was shown that the best results on the music relevant dataset is clearly improved by the correct initialization of the pheromone and heuristic values in an ant colony. Furthermore, it would be interesting to analyze, how the other parameter influence the results, e.g. the trade-off between exploitation and exploration under the control of the hyper-parameter, in this study $q_0 = 0.6$; or how the hyper-parameter μ changes the final results, if each ant has a bigger capacity for the feature subset, in this study $\mu = 0.35$. Some conclusions about how to select hyper-paremeters like β, ρ for other applications, e.g., for the well-known traveling salesman problem are drawn by [27]. Additionally, the optimal population size of ants in a colony in the feature selection process was discussed in [30]. How the other parameters are initialized in selecting the features, should also be investigated in the future.

Knowledge Base: There are music-related named entities that occur in German texts but there is no corresponding entry in databases like dbpedia or Musicbrainz for those entities. Hence, the availability of music-related entities in such databases plays a crucial role in expanding the knowledge base in order to assign correct music roles. Here, suitable fall-back strategies should be considered in future work.

Polysemy & Synonymy: During the experiments it has been noticed that no text representation captures phenomena like polysemy or synonymy of words. In German, the word *spielen* (play) has the following syntactic and semantic environment: “Orgel spielen” (playing organ music) and “gegen eine Mannschaft spielen” (playing against a team). For the former one, it is certain that it is music relevant, the latter implies a sport event. However, without further contextual information, even the first one might not be recognized as music relevant. Similarly, the problem of synonymy deserves much more attention in future work, e.g., “... in der Gosecker Schloss-Krypta klangen wieder kristallene Engel...” (in the Gosecker castle-crypt crystal angels sounded again). In this context, the word *klangen* means to play music. Those words are quite often used in public announcement in Germany. Thus, it could be a meaningful feature to decide whether a text is music relevant or not.

ACKNOWLEDGMENT

The authors would like to thank the deecoob GmbH for acting as experts during the creation of the gold standard and providing data. The project was funded by the German federal ministry for economics and energy.

REFERENCES

[1] J. Xi, M. Spranger, H. Siewerts, and D. Labudde, “Towards an automated system for music event detection,” in SEMAPRO 2018, The Twelfth International Conference on Advances in Semantic Processing. ThinkMind, 2018, pp. 22–27.

[2] GEMA, “Geschäftsbericht mit Transparenzbericht 2017,” https://www.gema.de/fileadmin/user_upload/Gema/geschaeftsberichte/GEMA_Geschaeftsbericht2017.pdf, 2018 [retrieved: September, 2018].

[3] V. Gupta and G. S. Lehal, “A Survey of Text Mining Techniques and Applications,” Journal of Emerging Technologies in Web Intelligence, vol. 1, no. 1, 2009, pp. 60–76. [Online]. Available: www.alerts.yahoo.com

[4] M. Spranger, H. Siewerts, J. Hampl, F. Heinke, and D. Labudde, “SoNA: A Knowledge-based Social Network Analysis Framework for Predictive Policing,” International Journal On Advances in Intelligent Systems, vol. 10, no. 3 & 4, 2017, pp. 147–156.

[5] K. Giridhar and A. James, “Text Classification and Named Entities for New Event Detection,” in Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 2004, pp. 297–304.

[6] A. Edouard, “Event detection and analysis on short text messages,” Ph.D. dissertation, Université Côte D’Azur, 2017, https://hal.inria.fr/tel-01680769/document [retrieved: September, 2018].

[7] M. Ikonomakis, S. Kotsiantis, and V. Tampakas, “Text Classification Using Machine Learning Techniques,” WSEAS Transactions on Computers, vol. 4, 2006, pp. 966–974.

[8] R. Bellman, “The Theory of Dynamic Programming,” Bulletin of the American Mathematical Society, vol. 60, no. 6, 1954, pp. 503–515.

[9] Y. Yang, T. Pierce, and J. Carbonell, “A Study of Retrospective and On-line Event Detection,” in Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, 1998, pp. 28–36.

[10] Y. Raimond and S. Abdallah, “The Event Ontology,” http://motools.sourceforge.net/event/event.html, 2007 [retrieved: July, 2019].

[11] J. Allan, Topic detection and tracking: event-based information organization, J. Allan, Ed., 2002.

[12] M. Cordeiro and J. Gama, “Online social networks event detection: A survey,” in Solving Large Scale Learning Tasks. Challenges and Algorithms, S. Michaelis, N. Piatkowski, and M. Stolpe, Eds., vol. 9580. Springer Verlag, 2016, pp. 1–41.

[13] T. Sakaki, M. Okazaki, and Y. Matsuo, “Earthquake shakes twitter users: Real-time event detection by social sensors,” in Proceedings of the 19th International Conference on World Wide Web. ACM, 2010, pp. 851–860.

[14] E. Lefever and V. Hoste, “A classification-based approach to economic event detection in dutch news text,” in Proceedings of the Tenth International Conference on Language Resources and Evaluation. European Language Resources Association (ELRA), 2016, pp. 330–335.

[15] G. Jacobs, E. Lefever, and V. Hoste, “Economic event detection in company-specific news text,” in Proceedings of the First Workshop on Economics and Natural Language Processing. Association for Computational Linguistics, 2018, pp. 1–10.

[16] F. Hogenboom, F. Frasincar, U. Kaymak, and F. de Jong, “An overview of event extraction from text,” in CEUR Workshop Proceedings, vol. 779, 2011, pp. 48–57.

[17] E. Benson, A. Haghighi, and R. Barzilay, “Event discovery in social media feeds,” in Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1, ser. HLT ’11. Stroudsburg, PA, USA: Association for Computational Linguistics, 2011, pp. 389–398.

[18] D. Nadeau and S. Sekine, “A survey of named entity recognition and classification,” Linguisticae Investigationes, vol. 30, 2007, pp. 3–26.

[19] M. K. Stefan Andelic, “Text classification based on named entities,” 2017, pp. 23–28.

[20] Y. Gui, Z. Gao, R. Li, and X. Yang, “Hierarchical text classification for news articles based-on named entities,” Advanced Data Mining and Applications, vol. 7713, 2012, pp. 318–329.

[21] M. Spranger and D. Labudde, “Towards Establishing an Expert System for Forensic Text Analysis,” International Journal on Advances in Intelligent Systems, vol. 7, no. 1/2, 2014, pp. 247–256.

[22] Y. Yang and J. O. Pedersen, “A comparative study on feature selection in text categorization,” in Proceedings of the Fourteenth International Conference on Machine Learning, 1997, pp. 412–420.

[23] J. Tang, S. Alelyani, and H. Liu, "Feature selection for classification: A review," Data classification: Algorithms and applications, 2014, p. 37.

[24] H. Liu and H. Motoda. Boca Raton, Florida: Chapman and Hall/CRC, Oktober 2007, ch. 2,13.

[25] M. Dorigo, "Ant Colony System: A Cooperative Learning Approach to the Traveling Salesman Problem," IEEE Transactions on Evolutionary Computation, vol. 1, April 1997, pp. 53–66.

[26] D. J. T. Sumpter, Collective Animal Behavior. Princeton (New Jersey): Princeton University Press, 2010, ch. 1,5.

[27] M. Dorigo and T. Stützle, Ant Colony Optimization. Cambridge, Massachusetts. London, England: The MIT Press, 2004, ch. 1,3.

[28] R. K. Sivagaminathan and S. Ramakrishnan, "A hybrid approach for feature subset selection using neural networks and ant colony optimization," Expert Syst. Appl., vol. 33, 2007, pp. 49–60.

[29] D. Martens, M. De Backer, R. Haesen, J. Vanthienen, M. Snoeck, and B. Baesens, "Classification With Ant Colony Optimization," IEEE Transactions on Evolutionary Computation, vol. 11, no. 5, 2007, pp. 651–665.

[30] M. H. Aghdam, N. Ghasem-Aghaee, and M. E. Basiri, "Text feature selection using ant colony optimization," Expert Systems with Applications, vol. 36, no. 3, April 2009, pp. 6843–6853.

[31] E. Saraç and S. A. Özel, "An Ant Colony Optimization Based Feature Selection for Web Page Classification," The Scientific World Journal, vol. 2014, 2014.

[32] M. M. Kabir, M. Shahjahan, and K. Murase, "An Efficient Feature Selection Using Ant Colony Optimization Algorithm," Neural Information Processing, vol. 36, 2009, pp. 242–252.

[33] M. M. Kabir and K. Murase, "A Wew Wrapper Feature Selection Approach Ssing Neural Network," Neurocomputing, vol. 36, 2010, pp. 3273–3283.

[34] F. Pedregosa et al., "Scikit-learn: Machine learning in python," J. Mach. Learn. Res., vol. 12, Nov. 2011, pp. 2825–2830. [Online]. Available: http://dl.acm.org/citation.cfm?id=1953048.2078195

[35] M. F. Porter, "An algorithm for suffix stripping," Program, vol. 14, no. 3, 1980, pp. 130–137.

[36] H. Schmid and H. Schmid, "Improvements in part-of-speech tagging with an application to german," Proceedings of the ACL SIGDAT-Workshop, 1995, pp. 47–50.

[37] J. R. Finkel, T. Grenager, and C. Manning, "Incorporating non-local information into information extraction systems by gibbs sampling," in Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics. ACM, 2005, pp. 363–370.

[38] J. Xi, "Music Classification Gold Standard Datasets," https://github.com/fosil-mw/music_classification_data, 2018 [retrieved: July, 2019].

[39] Q. Yang and S. J. Pan, "A Survey on Transfer Learning," IEEE Transactions on Knowledge Data Engineering, vol. 22, 2009, pp. 1345–1359.

[40] Y. Bengio, R. Ducharme, P. Vincent, and C. Janvin, "A neural probabilistic language model," Journal of Machine Learning Research, vol. 3, 2003, pp. 1137–1155.

[41] D. K. Xiao Li and C. X. Ling, "Active learning for hierarchical text classification," Tan PN., Chawla S., Ho C.K., Bailey J. (eds) Advances in Knowledge Discovery and Data Mining, vol. 7301, 2012, pp. 14–25.

[42] A. McCallum, R. Rosenfeld, T. M. Mitchell, and A. Y. Ng, "Improving text classification by shrinkage in a hierarchy of classes," in Proceedings of the Fifteenth International Conference on Machine Learning, 1997, pp. 359–367.

Managing Process Change and Standardisation in ERP Projects: An Assessment of the SAP Template Approach

Markus Grube

Barum, Germany
e-mail: research@markus-grube.de

Martin Wynn
School of Computing and Engineering
University of Gloucestershire
Cheltenham, UK
e-mail: MWynn@glos.ac.uk

***Abstract* — The use of templates for process alignment in ERP projects is an established practice and many companies have used this approach when implementing the SAP ERP software system. This research paper deals with the question of how companies use a template in a business process management environment and, more specifically, what happens when there are requests for changes to the template. Based on a literature review and personal interviews, this article assesses the commitments that manufacturing companies make regarding template implementation within SAP business process management. It suggests that "Defend the template!" may be an appropriate overarching goal.**

Keywords — SAP; ERP; Enterprise Resource Planning; BPM; Business Process Management; Template; Process change; Process standardisation; People capabilities.

I. Introduction

This article builds upon previously published research [1][2] which describe that Enterprise Resource Planning (ERP) software packages have been widely deployed since their emergence in the late 1980s, and on the role of Business Process Management (BPM) maturity models and ERP systems highlighted the significance and value of templates in aligning or remodelling business processes in such projects. Foth [3] defines a template as a unified set-up containing the necessary documentation and programmes for the construction of a new system. In addition to facilitating improved business processes, the use of a template offers numerous other potential advantages, including a reduction in implementation costs and required manpower, and an increased level of data consistency. This article examines how templates can be used to manage and improve business processes, and also assesses the importance of people skills and competencies in ERP systems projects. Following this introductory section, Section II discusses the relevant literature and develops the research questions for this paper. Section III then describes the research methodology and Section IV provides a summary of the research findings. Section V addresses the research questions and, finally, Section VI presents the conclusion and suggestions for possible future research.

II. Literature Review

An ERP system is generally seen as a software package which automates and integrates business processes, shares common data, and produces and accesses information in a real-time environment. ERP software can also be implemented in stages and therefore, be used to connect previously isolated IT systems and functional departments within a company. ERP is also viewed by some researchers [4][5][6] as a fundamental method for achieving best practice within business operations – the implementation of the ERP package requiring the application of certain disciplines within main business processes. According to Turban et al. [7], ERP not only provides business discipline, but it also allows the alignment of IT deployment with overall business strategy and business goals. Implementing ERP may thus also require a change in core processes, often termed business process reengineering or "BPR" [8]. The globalisation of business activities has reinforced the need to structure and standardise business processes across the enterprise and to adhere to corporate rules and procedures.

In general, "the main goal of process standardisation is the development of one standard or best-practice process to be used as a template for all instances of the process" [9]. Gavidia [10] asserts that it is common practice to develop a template as the base configuration for all areas and locations of a company to increase compliance and data consistency. This new template will then have a consistent set of reports and will be adapted to the existing business processes within the company.

There are a number of third party products that focus on process design and management. Metastorm, for example, uses a "BPM Designer" by which "you design the actual process, define business rules, set roles and associate forms. The main components of the process are where you start from, the steps or 'stages' that will be traversed, and the endpoints" [11]. ERP products like SAP usually offer "best-practice" templates for BPM in different industries and business processes. However, for various reasons, the suggested best practice is often found not to work well, and thus many companies define their own BPM templates to meet their specific needs [12]. In addition, it may be necessary to distinguish between a global template and local requirements and to define mechanisms that address such differences [13].

Some companies do not adopt a template approach at all, preferring instead to map, document, analyse and improve their processes using in-house methods, even when implementing ERP systems. This is particularly the case in small and medium-sized enterprises (SME), where processes may have developed and been customised over many years to meet specific ways of working or customer requirements and

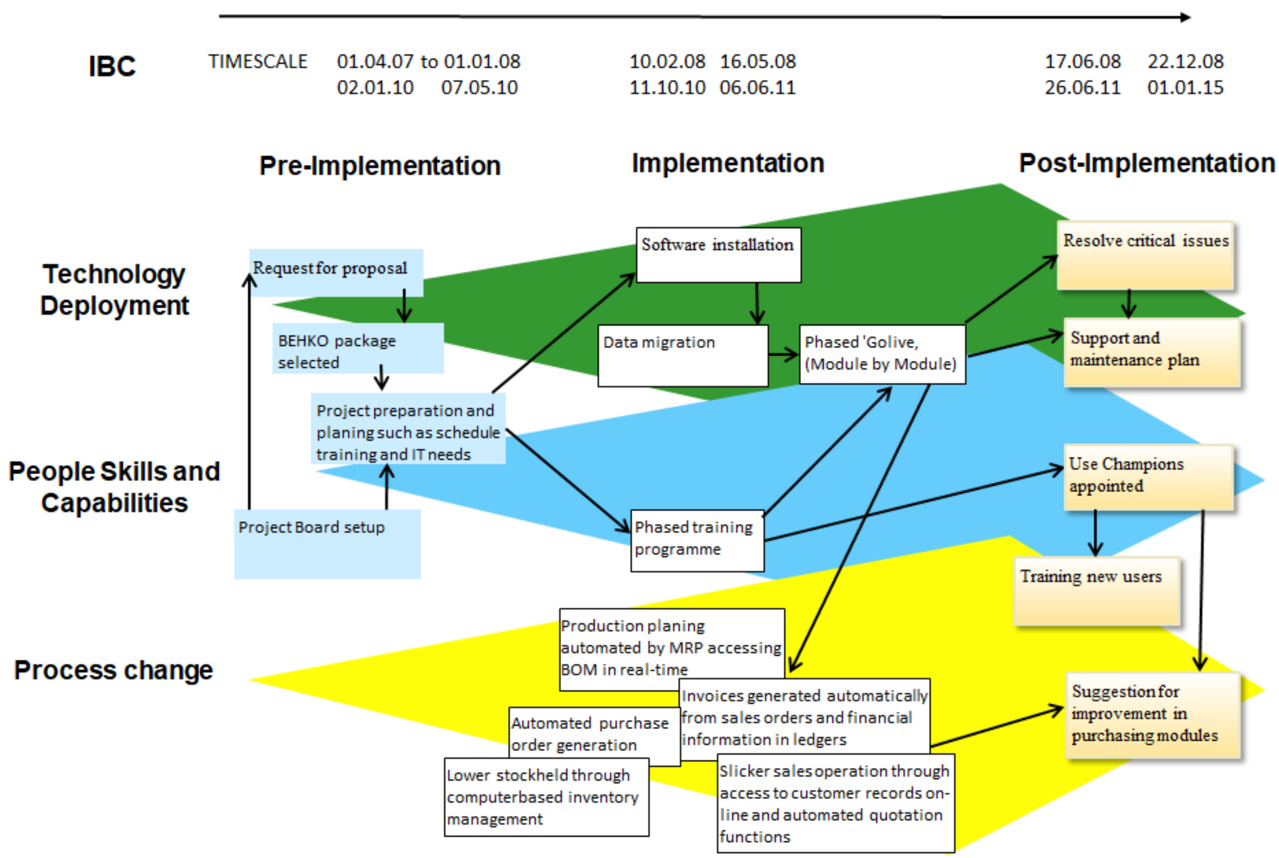


Figure 1. Implementation of the BEHKO ERP systems at the Isfahan Bus Company [15].

the template approach is thus difficult to implement [14]. This point is illustrated by a recent study by Rezaeian and Wynn [15] of ERP systems implementations in three Iranian manufacturing SMEs. The mapping of company processes provided the basis for selecting the best-fit ERP product available, but the modification of existing processes to fit the software was limited. Nevertheless, as the packages were implemented (sometimes module by module), small scale ad hoc process change allied to limited customisation of the software provided reasonably successful outcomes, notably at the Isfahan Bus Company (Figure 1). As Pajk and Kovačič [16] note, "these days ERP systems need to offer a lot of functionality in order to cope with a large number of business requirements. This functionality needs to be aligned with the business in order to create value for the organisation, confronting the organisation with the options of either configuring the enterprise system, the organisation, or a combination of both".

From an Information Technology (IT) perspective, it makes sense to standardise business processes and associated procedures and transactions. Costs are reduced if the organisation can adopt the business models inherent in an ERP system, rather than vice versa [17]. However, this approach is criticised in the literature. For example, Gavidia [10] reports that blanket implementation of a uniform template reflects a failure of management and risks damaging existing customer relationships. A process template can only be effective if it takes into account the operational and information requirements of all areas of a company.

Heeks [18] identifies the key aspects of change that provide a frame of reference for effecting the organisational transition associated with and required by, the implementation of an ERP system, and this adds an interesting perspective regarding the value of process templates. Heeks suggests that there are four aspects of an organisation that must change as it adopts major new information systems like an ERP product.

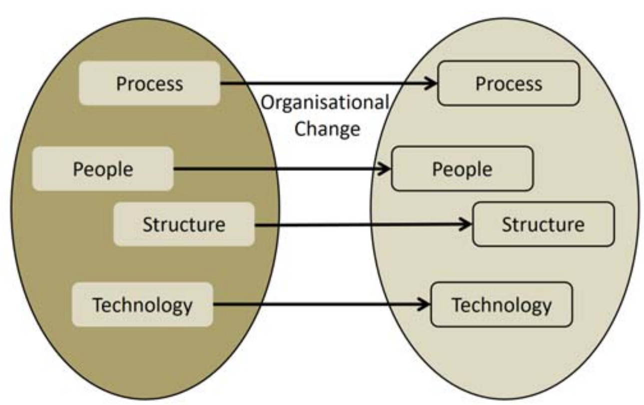


Figure 2. Design - Actuality Gap Model [18].

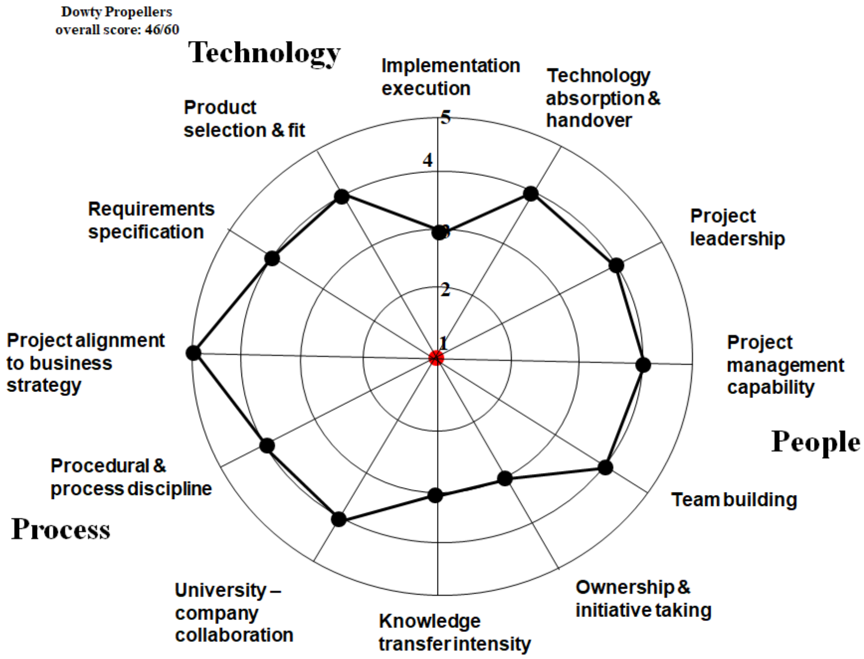


Figure 3. Change factors in the Dowty Propellers ERP implementation [20].

He sees these as constituting a "Design-Actuality" gap – i.e., a gap between where an organisation may currently be, and where it needs to be to successfully adopt major new systems. One of these four elements does indeed focus on Process, but the other three are of equal importance - People, Structure, and Technology (Figure 2). In similar vein, recently published research on technology transfer by Wynn [19][20] included a number of ERP projects in which a range of change factors were identified for successful implementation. These were grouped into three main dimensions of change, encompassing technology, process and people factors (Figure 3).

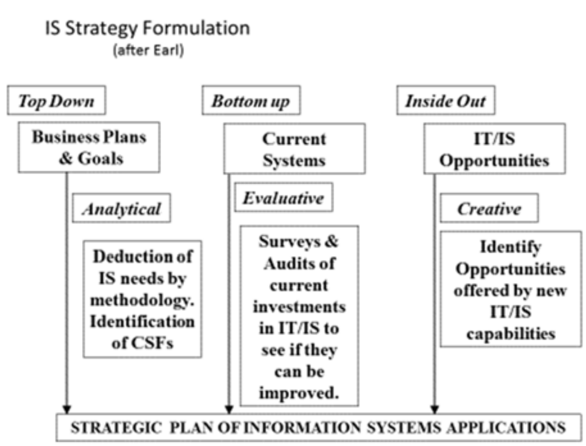


Figure 4. Earl's model of Information Systems strategy formulation [21].

Michael Earl's classic model of IS strategy formulation and implementation (Figure 4) is also of relevance to the debate regarding process templates. Earl's "three-pronged" model remains one of the standards in the analysis of information systems delivery. Although many see his three possible strategies as alternatives, Earl, in fact, argued that all three can be used in combination and that different strategies suit different business contexts. This depends on a number of factors including the nature of the basic IT infrastructure, the structure of the organisation and the level of development of the organisation's main business systems.

The "top-down" approach is driven by a formal attempt to match IS investment with business needs. It identifies the business plans and goals and then applies an analytical approach to identify information systems requirements. Critical Success Factors (CSFs) can be used to establish information and systems needs. It ensures business managers are involved in IS strategy formulation, which is thus seen as "top-down". The "bottom-up" approach is an evaluation of the capabilities of information systems currently in place. Earl notes that "IS strategies are rarely developed from a green field site, but have to recognise the strengths and weaknesses of the current applications portfolio" [21]. This may establish that some of the existing systems could be better exploited for strategic advantage or be improved upon to produce "significant added value". The "inside-out" approach attempts to "identify opportunities afforded by IT/IS which may yield competitive advantage or may create strategic options" [21]. An IS strategy based on deployment of an ERP package is likely to "top-down" at the outset, based on a strategic vision for the company and a preparedness to change business processes in pursuit of a more efficient company utilising a state of the art software system. However, as such projects progress, this may be complemented by a "bottom-up" approach, which may lead to the retention of some pre-existing software and challenge adherence to the ERP vendor's process templates.

With these considerations in mind, this article addresses two research questions (RQs):

RQ1: To what extent should the process template be adhered to in ERP implementations (with a focus on SAP templates)?

RQ2: What measures need to be taken to ensure People Process and Technology elements are kept in balance when using a strict template approach to process and systems management?

III. Research Methodology

The research on which this article is based takes a post-positivist position and is based on the perspectives of Ryan [22] and Guba [23]. The goal of research from the post-positivist perspective is to generate new knowledge that other people can learn from, and on which they can base their own decisions [24]. Semi-structured and in-depth interviews were used to gather practical experience of companies that use SAP templates and BPM. These interviews were followed by an analysis of the interview material and here an abductive approach [25] was adopted. This is an explanatory study, as defined by Collis and Hussey [26], which uses interviews with experts who have already gained practical experience in the areas of BPM in ERP projects and the use of SAP templates. It is quite possible that the involvement of other experts would lead to slightly different results, but this is in accordance with the post-positivist research position, which accepts that the world is more complex than that which the opinions of some experts may suggest.

To address the research questions, this research combined two research methods with the aim of achieving a greater depth of understanding of the complex environment being studied [27]. Following a literature search and review, semi-structured interviews were conducted with experts in their field. Figure 5 depicts the overall research strategy. The time horizon for this research is a cross-sectional snapshot study [28]. The research analyses the current practical situation and evaluates the state of affairs at the time of the study [29].

In the context of the semi-structured interviews, an expert had to have experience in the areas of SAP and BPM. Practical experience of ten years or more in the mentioned areas was expected. Due to the small number of interviews, this research did not differentiate between which modules of SAP these experts used or were familiar with, or within which industrial sector they had worked, or currently did work. Respondents were selected with the aim of gaining as much expert knowledge as possible from practice. Using semi-structured interviews provided flexibility to scrutinise the understanding and explanation of the experts' opinions [30].

The qualitative data analysis and the comparison of the transcribed interviews were supported by using the software tool MAXQDA.

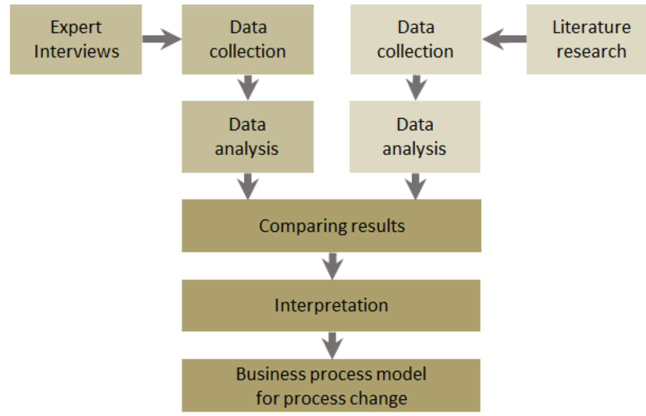


Figure 5. Research Strategy for the Template Study.

The coding functionality of the MAXQDA tool was used to analyse the experts' answers. The software was used to manually encode the qualitative data within the interviews and to obtain an overview of all the experts' answers based on the encodings. Due to the small number of interviews, the aim is not to reach an overall and general conclusion for all industries, organisations or SAP modules. However, the findings may well provide a platform and theoretical grounding for further research in this field of study.

Sixty-four people were identified as potential experts in the business process and SAP environment, drawn from the authors' personal and business contacts in Germany, Austria, and Switzerland. These people were then invited for an interview. Many of those asked for an interview declined, citing their reluctance to pass on internal and confidential company data. Therefore, from this initial pool, eleven people confirmed that they were willing to be interviewed. The findings from these face-to-face interviews are discussed below. All statements and insights described here are the personal opinions of the experts. The total duration of the interviews was in excess of 14 hours in total and only excerpts are presented here.

IV. Findings

This section reviews the experts' perspectives on the SAP templates and business processes. It discusses the nature of templates, the implications of "localization" of processes, the trade offs between flexibility and standardization, the role of human factors, and the need for cross-company buy-in to process change.

Interview feedback suggested that the SAP ERP system offers a variety of "out-of-the-box" business processes for the general SAP system, as well as those for special business sectors, such as for the real estate or the pharmaceutical industries. With the basic process templates, a company can cover almost all of its standard processes, which can provide a solid basis for assessing the (possibly more idiosyncratic) daily business activities. This can help a company to standardise its processes and rapidly implement changes accordingly. For example, sales processes are similar in different companies and across different industries. It is helpful to use a standard SAP template for the best-practice approach as a starting point that provides a step-by-step illustration of the best-practice process sequence. In summary, these SAP templates can show and illustrate process details for general valid standard processes.

In such a standard SAP process template is embodied considerable industry expertise that has been accumulated from various industries and many projects. However, this does not mean that this template is optimal for any particular company, nor that the company's own process steps, which were developed and implemented in-house – and may deviate from the corresponding SAP process specifications - are not appropriate for that company. The experts pointed out that an approach that slavishly follows the given SAP processes is usually only used if a company has no existing processes or IT infrastructure. The interviewees pointed out that a single process initially has nothing to do with IT, but IT does support the later execution of the processes. As noted above, SAP offers the "out of the box" or "best-practice" processes which illustrate how IT supports business process behaviour.

Normally, a company has already established its own process flows, based on its perceived business needs and past experience. In these circumstances, such a company would not fully use the SAP "best-practice" standard template, but rather develop its own process template based on the SAP standard. This template must then have a certain degree of flexibility. A tax-paying process in Germany is different from one in, for example, Brazil or India. The basic process templates must be flexible enough to be used globally in different business and cultural environments, where there will be a certain amount of local variation in rules, regulations and working practices. Local adaptation must only be set up where a process, or a process step, is genuinely local and cannot be supported by a global SAP template. However, this means that the development of a company's own solution is divergent from the global template, creating a "shadow" process, which is only used for local requirements. The disadvantage of this route is that process change then becomes more difficult because the adaptation is known only locally, on the level of one company or business unit.

In light of the above, the experts emphasised that process templates should be as flexible as possible from the beginning of the set-up. They should be scalable and applicable to specific local needs. In the interviews, the experts identified this as the biggest challenge for a process template. A key point for consideration is whether new requirements that emerge can still be incorporated into the applied process template, and how quickly subsequent changes can be implemented company-wide. A process template must allow for a certain amount of agility to obviate the need for shadow process development.

In addition to the previous statements, the interviewees also reported that a process change is often triggered by humans and their adaptation needs. The problems linked to the introduction of such a process change and the subsequent use of process templates are often related to human behaviours. The human being is, therefore, a big driver but also a big preventer. Nearly every employee must be confident that such a template implementation really supports the company's needs and requirements. Some employees will always come up with reasons why their company is absolutely special and cannot use the standard template, and thus needs to keep with their own processes and IT programming. It is thus possible that another significant problem within implementing a process template is the human factor and the fact that some template descriptions and statements could be misinterpreted. As discussed by Hall [31], people can play an active role in the decoding and interpretation of a template. Each employee applies a template in the manner in which this use makes sense to him or her personally, even if this differs somewhat from the original intention of the template designer.

One of the experts emphasised the importance of the human factor in template implementation. This encompasses not only the prior knowledge of each employee but also the interpersonal behaviour between departments or among team members. It may thus be appropriate for the human resources department to be involved in process change and template introductions to oversee human relationship issues and potential conflicts. Not only must the technical aspects of template introduction be addressed, but personal relationships may also need to be clarified.

Regarding the use of standard processes, one expert raised the issue that business consultancies, when advising client companies, frequently use standard SAP templates without adequate assessment of the individual processes within the client company. This means that standard templates are established without optimising them to "best-fit" existing processes and special behaviours of each company. On the other hand, there are clear benefits in process standardisation through the introduction of a process template. This means, for example, that in a multi-site manufacturing company, each site would not be allowed to conduct its business according to a business process it wants and likes, but that a company-wide standard is set up and must be followed. The more complex and individualised the established processes are, the harder it is to run the business on one new company-wide process template. Uniform processes can create synergies and clarity, and a process template should help to minimise fragmentation and process variations in the company. The experts also maintained that there is a need for every company to analyse their processes individually, in order to uncover improvement potential. However, an organisation must also realise that non-individualised solutions can also be effective. Some of the experts suggested that nearly 80% of processes are very similar in industrial companies, and only a very few processes need an individual solution.

One expert gave the view that localisation of processes is a challenge because it can only be developed in collaboration with the local business departments. He highlighted the fact that the business must then maintain and update such localised processes, their templates and documentation because only the local team know how they really work. That said, another expert opined that the overall aim is to provide a standardised process template that is not subject to localisation and could be protected against further variation and customisation.

In this context, one interviewee mentioned the saying "Defend the Template", which is apparently current in some consultancy circles. The implication here is that a business should think very carefully about the repercussions of changing processes away from the template guidelines.

Another expert also declared that a template should be a "leader" for all processes, with functional processes introduced via a company-wide, valid template. The question is then no longer how the individual divisions or business units within an organisation carry out a task; instead, one "valid" process is established throughout the whole company. A process in the enterprise can then only be arranged based on these template specifications; nevertheless, there exist different variations, which are offered by a catalogue structure within the template. The potential benefit of minimising process diversity is that this may result in a significant reduction of costs and overheads. The goal of a process template must be to offer stable and mature process solutions for a company. A general default template must be defended against any further fragmentation or diversification within an organisation. The experts observed that it requires a lot of time and money to realise individual solutions, and define and develop unique processes involving specific coding.

These processes are then rolled out company-wide, and it is ideal to undertake a process without local on-site adaptation because these local adaptations cost in term of time, money, and manpower. A hurdle to avoid is being too detailed in the description of a process template. The more detailed the description of a process, the more likely it is to be a description of localised or specific particularities. The objective must be to develop an enterprise-wide process template that works equally well in a five-person facility as it does in a business unit with 800 employees in another country.

One of the interviewees stated that the use of SAP processes can be the catalyst for the introduction and assessment of standard business processes. Overall the findings from the interviews suggested there are then two ways forward. On the one hand, these standard processes can be increasingly embedded in company operations in a specific, individualised manner. On the other hand, companies may come to realise that not every process has to be individualised and that specific and individualised processes are rarely the best solution for a company overall. If a large ERP system like SAP has to be adapted, this means that human IT support will always be required. A system like SAP is so complex that changes in the SAP system cannot be adopted by everyone. Most of the time, a company relies on experts for these changes. The consequent additional time and cost must be taken into account when pursuing this option.

Using fewer individualised solutions and more standard process templates means that process changes can be implemented more quickly. This can also result in lower costs. Every process change, every system upgrade and change release - all carry the possibility of failure, and the risk of serious operational implications is reduced if there are fewer variations in the template catalogue that need to be taken into account. "Defend the template" aims to encourage the optimal use of process templates. With the resultant harmonisation of processes and operational activities, businesses should and can be convinced that the use of standard processes is not a negative approach.

But even if a standard template is to be adhered to, there must be scope to amend it or customise it when necessary; however, such a change should never be decided by one business unit alone. Such decisions should be taken by a cross-company body with overall process responsibility, consisting of, for example, process owners and process managers. Such a cross-organisational body can best assess the wider XXX implications and decide whether, and to what extent, processes can be changed or whether a business unit may use a local adaptation.

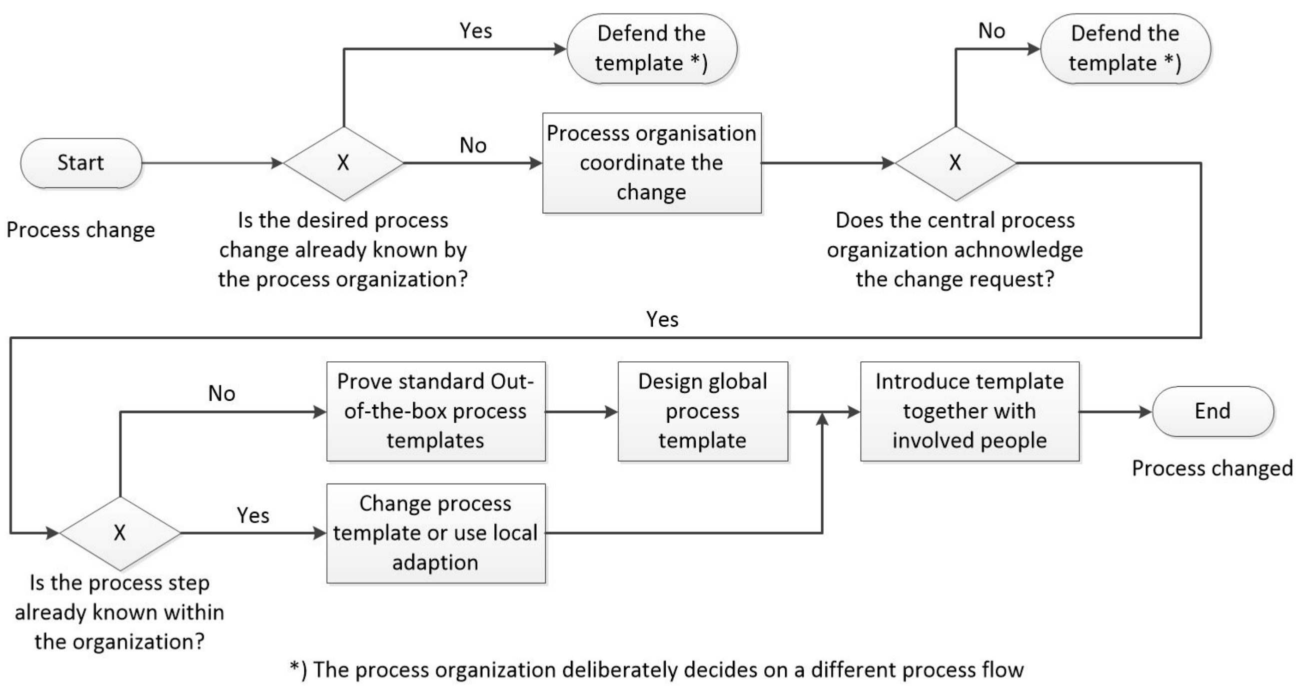


Figure 6. A simple model for process change and defence of the template.

V. Analysis

A combination of evidence from both the literature review and the expert interviews provides the basis for addressing both RQs.

RQ1: To what extent should the process template be adhered to in ERP implementations?

On balance, both the literature review and the interviews suggested that a lack of standardisation in the use of process templates should be avoided it at all possible, as this will lead to too much diversification to the detriment of the organisation overall. This manifests itself in the sub-optimal integration of processes and in unnecessary overheads in effecting future system changes and upgrades.

In the specific context of the SAP system, companies need to be aware that harmonising and optimising global business processes in, for example, a large multi-national company is an extensive project that can last for years. If variations have been introduced into a template, it will be more difficult and time-consuming to maintain this template and to manage every future change process within it. Companies should search for the lowest common denominator and defend the existing template as far as they can.

Business process management tools can be used to capture and document processes, allowing a company to visualise, sort and structure its processes and their management. This may lead to a degree of structural change and reorganisation (as noted in Heeks' model [18]); it often becomes evident that processes in different business units are similar. A process template must be built from these documented and structured processes, including all recognised process, flow, and, if appropriate, excluding certain process features and declaring them invalid.

For the practical introduction and implementation of the newly developed processes, the process team then has the task of establishing and defending these process templates and processes, even if they encounter resistance. Just because a process has been carried out in a company for years in a particular way, does not mean that this approach was the optimal solution for the business. However, when such a traditional and established process is changed, it is important to clarify why the new solution is a more effective approach and to present its benefits to all parties involved. Especially in such a case, the new process template must be defended, even if there may be strong resistance to it.

With such an extensive application of process implementation and the use of a template, this implementation has a lot to do with the human factor. It is important to address those involved team members directly and to integrate them into the template introduction process. Participants need to understand why a template is to be introduced and how it will be used.

As discussed above, some experts reported that business consultants like to introduce an SAP template without carrying out an individual process examination of the company. This kind of process template usage is to be avoided and every template introduction in the company must be checked for its usability before it is used.

The "defend the template" principle can be viewed as the essential starting point in a procedure for investigating process change requests. The difficulty is estimating the extent to which a template should be defended, as there can be legitimate doubts that a particular process template is an optimal solution for a company. In this case, a process owner or a process organisation must decide either to change the process or to make a local adjustment. Decisions in such exceptional cases should not be made by a business unit itself, but by a superior process organisation, which is responsible for the process and thus has to approve a change or local adaptation.

Customization of standard processes is sometimes unavoidable because an identical, single global process can never be appropriate for all scenarios - for example, when country-specific legislation has to be adhered to. This then contradicts the "defend the template" approach. In these cases, the challenge is to introduce country-specific modifications but to use the standard processes as much as possible.

From an analysis of the literature and the expert interviews, a procedure for changing a process template has been developed to provide an easy-to-use model to defend possible process changes. Such a procedure could consist of the steps shown in Figure 6. This figure contains two important gateways for a process change. First, the process organisation must be informed about the desired process change. If the process organisation already knows about the desired process change and has opted for another solution, then this alternative solution must be defended. The same applies to the second gateway. If the central process organisation does not accept a new change request, this decision must be defended by the process organisation against any resistance that may arise.

The process steps depicted in Figure 6 demonstrate that a template can be defended in several different locations within a process change. The advantage of this rigorous approach is that an escalation in local adaptions and shadow processes can be avoided.

In the context of the model previously put forward by Grube [1], care should be taken to use standard out-of-the-box process templates wherever possible, and to keep process extensions simple. Diversification from standard process functionality will create increased effort later on. Nevertheless, customer-specific process development must always be explicitly analysed when system changes are made.

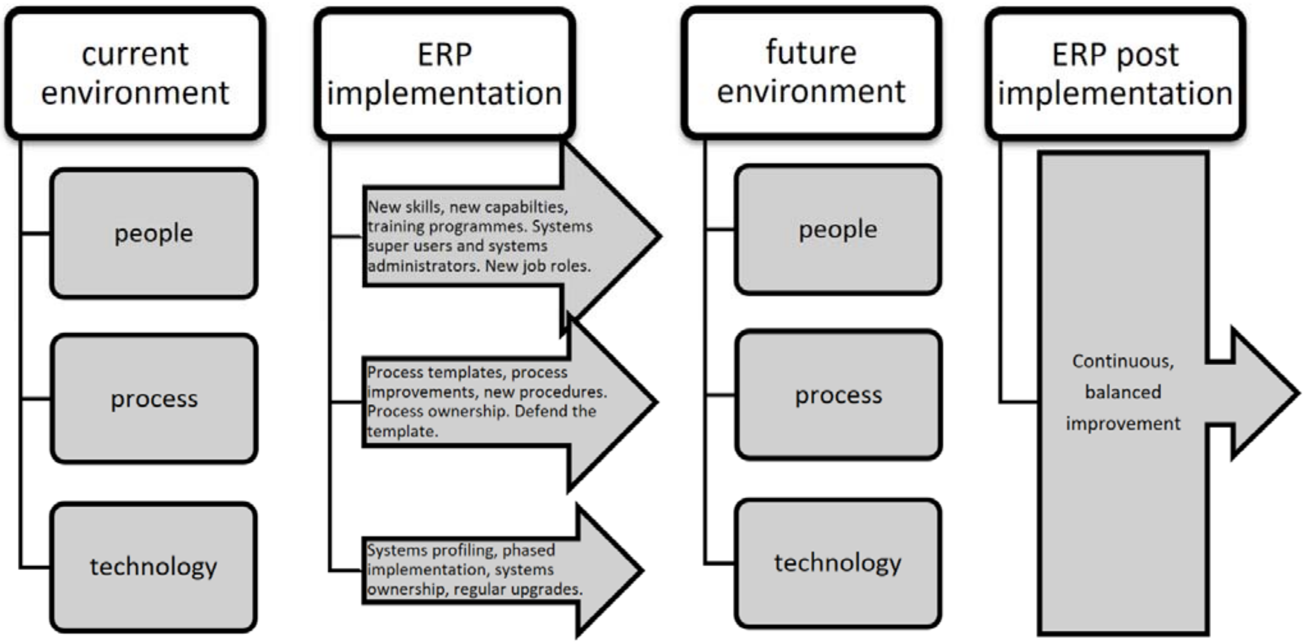


Figure 7. Continuous, balanced improvement in People, Process and Technology dimensions for ERP projects.

RQ2: What measures need to be taken to ensure people aspects are given appropriate focus when using a process template approach in an ERP implementation?

In addition to significant changes to processes, an ERP implementation may well require a balanced introduction of new staff skills and technology capabilities, as well as changes to organisational structure (Figure 7). For successful template implementation, all employees and appropriate departmental functions must be included. Only then can productive implementation be carried out successfully, with employees understanding the benefits and accepting the changes. A company should also consider that a template rollout requires a long-term implementation phase, which can take 2–5 years, and may require a budget of $10 million or more in a large international company [13].

The enforcement of a template is a practical difficulty in the process rollout. Often, this causes implementation delays and project and business goals can be impacted. A global template must be approved and accepted by all affected business units and employees, with appropriate briefings, workshops and training programmes. This requires discipline from all employees involved in the implementation of the template, because things may be repeated during the rollout, or there may be resistance against the implementation [32].

When implementing a new process template or defending a template, the responsible process department should not be afraid to bring the human resources department into the process. When implementing and defending process standards, HR departments can offer communicative support that an IT professional might have mediated differently. In the event of significant change, the people affected by the change must be considered. It is known from various sources that most people dislike any form of change. In this context, a well-known quote from an unknown author is: "The only person who doesn't resist change is a baby with a wet diaper" [33]. As the experts confirmed, it is imperative that people implement changes; but to do this successfully, the people involved must be convinced of the value of template application. A literature search would yield a large number of articles about the successful implementation of change. The three-step change model of Lewin is of relevance here as an illustrative example. This model represents a simple approach to change management and divides the change process into the three steps or phases - unfreezing, changing and refreezing [34]. The first phase helps to introduce and prepare employees because many people are against changes. Everyone involved in a change must first be prepared for this change. Lewin calls this phase unfreezing. The second phase is where the step-change is carried out and people learn the new behaviour. The final phase involves the refreezing step in which the change is stabilised and embedded (Figure 8).

Implementing ERP systems may entail significant changes to job briefs and roles, often involving new working practices. Employees will be required to change, and, more importantly, they should feel involved in that change. It is therefore crucial that employees understand the rationale for the new ERP system, the process changes required and the significance and value of the process template, and feel shared ownership of the new system, the new processes and the project as a whole. There are a number of tools and methods for charting the change in the people capabilities as an ERP system is implemented. For example, the People Capability Maturity Model (PCMM) may be of value in monitoring changes in workforce skills and capabilities. This model consists of five maturity levels - Initial, Managed, Defined, Predictable and Optimising – that can be applied at the process level to help gauge advances in people skills and coordinate them with related process and technology change. This can lead to a program of continuous workforce development within which training programmes can be embedded and cross-referenced with process improvement and technology deployment [35].

VI. CONCLUSION

The findings from the literature review and the expert interviews suggest that a suitable compromise must be found between standardisation, (i.e., adherence to the template) and flexibility to adapt to local requirements. The practical experience of the experts interviewed in this research demonstrates that the "defend the template" principle can act as a useful starting point for process management and alignment. Processes of all units within a business should be considered before a process template is created. Despite this preliminary work, a process template may have to be adapted later, but this adaptation must be coordinated by the process organisation, and must not be decided locally by a single business unit. If a suggested adaptation is not supported by the process organisation, the "defend the template" principle applies, as depicted in Figure 6.

The coordination of process change and technology deployment with people skills enhancement, new capabilities and appropriate training programmes are of paramount importance. These three dimensions of change should be managed in unison (Figure 7), and thus the "defend the template" principle could usefully be amended to "defend the template – but don't forget the people". In terms of Earl's model, a "bottom-up" approach can become of equal importance in implementation as a "top down" approach may have been in determining the systems strategy itself. Even as we move into an era of increased use of robotics, chatbots and artificial intelligence in many aspects of business life, people skills and capabilities remain a vital element of major systems projects.

This research clearly has its limitations. Only a few experts were interviewed, and it is quite possible that the inclusion of other experts, different industries or other ERP products would have led to different or additional results. Future research should closely examine the arguments put forward in this article, with a view to encompassing different business environments and ERP products. In addition, further research with more participants could significantly increase responses and lead to more diverse outcomes and analysis. This could, for example, include an online questionnaire to validate the steps shown in Figure 6, allowing the collection of a larger amount of data from more participants from different business and consultancy backgrounds and with different knowledge and experience of ERP systems and BPM products and practices.

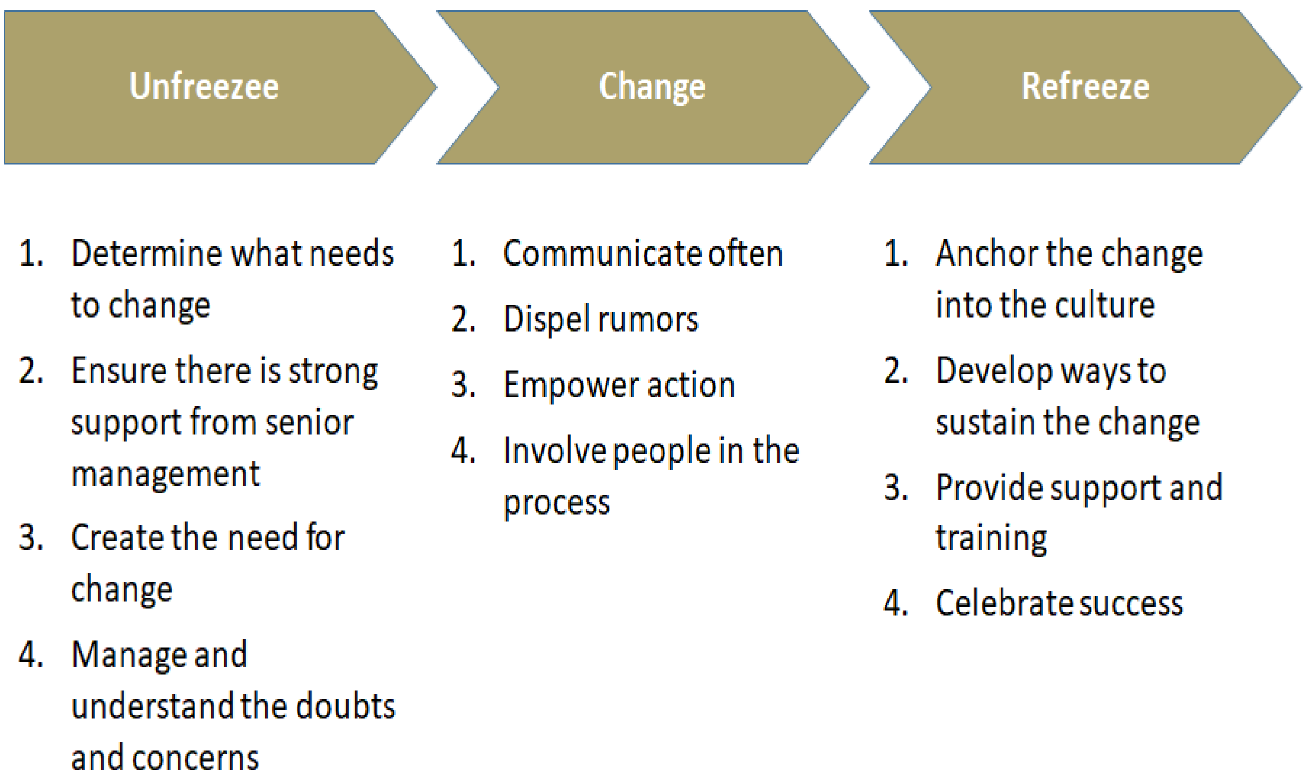


Figure 8. Lewin's three-step model for change management [36].

REFERENCES

[1] M. Grube, "Defend the Template – Supporting the Standardisation and Harmonisation of Process Change Programmes with SAP template", eKNOW, The Eleventh International Conference on Information, Process and Knowledge Management, February 2019, pp 20-25. ISSN: 2308-4375. ISBN: 978-1-61208-690-3. Available at: http://thinkmind.org/index.php?view=article&articleid=eknow_2019_2_10_60034.

[2] M. Grube and M. Wynn, "The Impact of SAP on the Utilisation of Business Process Management (BPM) Maturity Models in ERP projects", eKNOW 2018, The Tenth International Conference on Information, Process, and Knowledge Management, March 25-29, 2018, Rome, Italy. Available at: http://eprints.glos.ac.uk/5478/3/The%20Impact%20of%20SAP%20on%20the%20Utilisation%20of%20Business%20Process%20Management.pdf

[3] E. Foth, "Excellent business processes with SAP: Practice of use in business groups", (Exzellente Geschäftsprozesse mit SAP: Praxis des Einsatzes in Unternehmensgruppen), Springer Verlag, Berlin, Germany, 2010.

[4] C. Soh and S. Sia, "An institutional perspective on sources of ERP package organization misalignments", J. Strategic Information Systems, Vol.13, No. 4, pp. 375-397, 2004.

[5] K. Boersma and S. Kingma, "Developing a cultural perspective on ERP", J. Process Management, Vol. 11, No, 2, pp.123- 136, 2005.

[6] M. Arif, D. Kulonda, J. Jones, and M. Proctor, "Enterprise information systems: technology first or process first? " J. Business Process Management, Vol. 11, No.1, pp. 5-21, 2005.

[7] E. Turban, E. McLean, J. Wetherbe, N. Bolloju, and R. Davison, "Information Technology for Management – Transforming Business in the Digital Economy", 3rd edition, John Wiley & Sons, New York, 2002.

[8] M. Hammer and J. Champny, "Re-engineering the Corporation: A Manifesto for Business Revolution", Harper Business, New York, 1993.

[9] H. Romero, R. Dijkman, P. Grefen, and A. van Weele, "A literature review in process harmonization: a conceptual framework", Beta Working Paper series 379, Eindhoven, 2012.

[10] J. V. Gavidia, "Impact of parent-subsidiary conflict on ERP implementation", Journal of Enterprise Information Management, Vol. 29, No. 1, pp. 97-117, 2016.

[11] Process Mapping Pty Ltd, "Introduction to Metastorm" BPM, 2011. Available at: http://www.processmapping.com.au/articles/introductiontometastormbpm/ArchitectureandConfiguration.html.

[12] T. Kobayashi, S. Onoda, and N. Komoda, "Workflow Business Template for Application Processes in Administration Department", Information Technology and Management, vol. 3, no. 43, DOI: 10.1023/A: 1013160725421, 2002.

[13] SAP Blogs, "Overview of Global SAP Template Rollout Programs", Available at: https://blogs.sap.com/2011/12/11/overview-of-global-sap-template-rollout-programs/, 2020.05.24, 2011.

[14] M. Wynn and M. Rezaeian, "ERP implementation in manufacturing SMEs: Lessons from the Knowledge Transfer Partnership scheme", In Impact: The Journal of Innovation Impact, Vol. 8 No.1, 2015. Available at: http://eprints.glos.ac.uk/2930/1/ERP%20implementation%20in%20manufacturing%20SMEs.pdf

[15] M. Rezaeian and M. Wynn, "The implementation of ERP systems in Iranian manufacturing SMEs", International Journal On Advances in Intelligent Systems, Vol. 9, Nos. 3&4, 2016, pp.600 – 614, ISSN: 1942-2679.

[16] D. Pajk and A. Kovačič, "Fit Gap Analysis - the role of Business Process Reference Models", Economic and Business Review, Vol. 15, No. 4, 2013, pp. 319-338.

[17] A. Richen, and A. Steinhorst, "Standardization or Harmonization? You need both", available from: https://www.bptrends.com/publicationfiles/11-05-ART-StandardizationorHarmonizationv-RickenSteinhorst.pdf, 2020.05.24, 2005.

[18] R. Heeks, "Information Systems and Developing Countries: Failure, Success and Local Improvisation", The Information Society, Vol. 18, pp. 101-112, 2002.

[19] M. Wynn, "Technology Transfer Projects in the UK: An Analysis of University-Industry Collaboration", International Journal of Knowledge Management Vol. 14, Issue 2, 2018, pp. 52-72. ISSN 1548-0666.

[20] M. Wynn, "University-Industry Technology Transfer in the UK: Emerging Research and Future Opportunities". Advances in Knowledge Acquisition, Transfer, and Management (AKATM) IGI-Global, Hershey, USA, 2018.

[21] M. Earl, "Management Strategies for Information Technology", Prentice Hall, Hemel Hempstead, 1989

[22] A. B. Ryan, "Post-Positivist Approaches to Research Researching and Writing your thesis: a guide for postgraduate students"; MACE, Maynooth Adult and Community Education, pp. 12-26, 2006.

[23] E. G. Guba, "The Paradigm Dialog"; Newbury Park, CA, Sage Publications, 1990.

[24] Z. O'Leary, "The Social Science Jargon Buster", London, UK, SAGE Publications Ltd, 2007.

[25] G. Thomas, "How to Do Your Case Study: A Guide for Students and Researchers", London, UK SAGE Publications, 2011.

[26] J. Collis and R. Hussey, "Business research: A practical guide for undergraduate & postgraduate students"; Basingstoke, UK, Palgrave Macmillan, 2009.

[27] R. K. Yin, "Case Study Research: Design and Methods", London, UK, SAGE Publications, 2009.

[28] M. Saunders, P. Lewis, and A. Thornhill, "Research methods for business students", New York NY, Prentice-Hall, 2009.

[29] R. Kumar, "Research Methodology: A Step-by-Step Guide for Beginners", London, UK, SAGE Publications, 2011.

[30] A. Bryman and E. Bell, "Business research methods", Oxford, UK, Oxford University Press, 2007.

[31] S. Hall, "Encoding and Decoding in the Television Discourse," University of Birmingham, 1973, Available at: https://www.birmingham.ac.uk/Documents/college-artslaw/history/cccs/stencilled-occasional-papers/1to8and11to24and38to48/SOP07.pdf, 2020.05.24.

[32] E-3 Magazin, "Template-based SAP rollouts", Available from: https://www.cbs-consulting.com/files/pdf/information/Templatebasierte-SAP-Rollouts_E3_092011.pdf, 2020.05.24, September 2011.

[33] Anonymous, New Castle (PA) News, "Fox endorses amendments," pg. 25, col. 8, 8 May 1967.

[34] K. Lewin, "Frontier in Group Dynamics: Concept, Method and Reality in Social Science; Social Equilibria and Social Change". In: Human Relations. Vol. 1, no. 1, pp. 5 – 41, ISSN 1573-9716, DOI:10.1177/001872674700100103, 1947.

[35] B. Curtis, W. Hefley, W. E. and S. Miller, A. People, "Capability Maturity Model (PCMM)", Version 2, Software Engineering Institute, 2001.

[36] Mind Tools Ltd, "Lewin's Change Management Model - Understanding the Three Stages of Change". Available at: https://www.mindtools.com/pages/article/newPPM_94.htm, 2020.05.24, 2016.

Guidelines for Participatory Design with People Living with Mild Acquired Cognitive Impairments

Klaudia Çarçani
Faculty of Computer Science
Østfold University College
Halden, Norway
email: klaudia.carcani@hiof.no

Harald Holone
Faculty of Computer Science
Østfold University College
Halden, Norway
email: harald.holone@hiof.no

***Abstract* –Mild Acquired Cognitive Impairment(s) (MACI) are called invisible impairments. This invisibility of symptoms makes that researchers often overlook people with MACI. This is the case in the field of Participatory Design (PD). In this paper, we investigate how to involve people with MACI in designing together digital solutions meant for them. Hence, how to involve people with MACI in Participatory Design. Considering the lack of literature in PD focused on MACI patients, we conducted a borderer investigation of the literature and derived a set of guidelines proposed by PD practitioners for involving people with cognitive impairments in PD. We have collected data on conducting PD with MACI patients from two empirical cases as part of two projects in a rehabilitation hospital in Norway, which offers specialized rehabilitation to people with MACI. We conducted 5 PD workshops with three different workshop outlines. Seventeen people with MACI participated. We present a detailed list of reflections-on-action for each workshop outline. We discuss the reflections with findings from the literature and conclude with a list of guidelines that researchers and designers should consider when involving people with MACI in conducting PD. The paper aims to attract the attention of PD practitioners to the MACI user group with the possibility of expanding the guidelines list in the future.**

Keywords-Participatory Design; Mild Acquired Cognitive Impairments; Guidelines.

I. INTRODUCTION

There is an increase in chronic diseases in our aging society, and technology is considered as one of the means to cope with the increasing number of people living with chronic conditions. A notable case is individuals with Acquired Brain Injury (ABI) affecting cognitive capacities. For this user group, the use of technology has become a fundamental part of their daily lives by providing a wide range of useful services and tools to use at home, work, or anywhere else. This paper expands on previously reported results on how to involve people with ABI in designing new technologies that can facilitate their everyday life challenges [1]. Intensive research is ongoing regarding technologies that support patients with moderate or severe cognitive impairments (i.e., [2][3][4]). However, less attention had been paid to people suffering from mild cognitive impairments after an Acquired Brain Injury (ABI), even though mild acquired cognitive impairments are a critical global public health problem and listed among the major causes of permanent impairments [5][6]. In this paper, the abbreviation "mild acquired cognitive impairment(s)" (MACI), coined from Eghdam, Scholl, Bartfai, and Koch [7], will be used to refer to these people. Note that this is not an official abbreviation for the clinical condition. MACIs are usually described as invisible impairments and might include problems with memory, attention, executive functioning, language, and fatigue. People suffering from MACIs typically have a very challenging daily life, given the invisible nature of the condition.

An essential factor for the design of new technologies that can assists MACI people is their involvement in the design of these solutions that will be used by them in the future. User participation constitutes the core of Participatory Design [8]. PD is that design approach that promotes the involvement in the design process of everyone that will be influenced by the newly designed technological solution. Thus, new technologies aiming to assist individuals who have had an ABI would, from a PD perspective, require people with MACIs to be involved in the design process. However, involving people that have cognition problems to envision future solutions or discuss abstract terms is challenging. PD researchers should consider how to promote involvement and participation. The research question we raise in this paper is *What should be taken into consideration when conducting Participatory Design with people with MACIs?*

PD has, in decades, been concerned with marginalized groups and research on how to involve people with dementia, aphasia, amnesia, cognitive impairments, elderly, etc. in the design of new solutions. While the case of people with MACI is different from these diseases, some symptoms are similar. Thus, a look into the existing literature and how other researchers have described guidelines for working with user groups similar to MACI is relevant for building preliminary knowledge. In this paper, we initially present a summary of guidelines for involving people with cognitive impairments in PD projects. The guidelines are from a systematic literature review conducted by the authors of this paper. The findings were cross-checked and refined with the findings of a literature review of conducting PD with people with dementia by Hendriks, Truyen, and Duval [9].

Further, we present two empirical cases of conducting PD with MACI people. We take a reflective practitioner analysis approach to derive a set of reflections for involving MACI people in PD. These reflections, which emerged from the empirical data from 5 workshops with a total of 17 MACI patients as participants, have been further cross-analyzed with the list of guidelines found from the literature. The cross-analysis is also influenced by the literature on MACI people, their symptoms, and advice for leaving with cognitive challenges from medical practices. We have concluded with a list of guidelines for conducting PD with MACI people.

We contribute by expanding the knowledge of conducting PD with a special user group like MACI people who represent a huge share of the society [6] and can benefit from technology to facilitate their daily life. Moreover, we contribute by bringing the attention of the PD researcher to this user group, which due to the "silent" characteristics of the disease, are usually forgotten or overlooked and left alone to fight a challenging daily life.

The next section gives a more detailed overview of the user group, the symptoms they have, and some suggestions from the medical perspective on how the MACI people can adapt their life to the cognitive challenges. This is followed by a section on what PD is and why it is relevant to conduct PD with MACI people. We present a detailed summary of guidelines deriving from previous research on conducting PD with people with cognition challenges as part of our background literature. Further, we describe our methods for collecting the data and analyzing it, followed by the empirical material from two projects done with MACI people. We describe three workshop outlines applied in 5 PD workshops in the two projects. A list of methodological reflections [10] follows each workshop outline. We further analyze and discuss the findings from the empirical cases with the data from the list of guidelines coming from the literature review. In conclusion, we present a set of guidelines for conducting PD with MACI people.

II. MILD ACQUIRED COGNITIVE IMPAIRMENTS AND THE IMPLICATIONS IN A PERSON'S LIFE

In this section, we first describe what it means for a person to live with MACIs. Further, we define what cognitive rehabilitation is and how this service is offered in a rehabilitation hospital in Norway, where we conducted our research.

A. *What does it mean to live with Mild Cognitive Impairments (MCI) after Acquired Brain Injury (ABI)?*

Cognition is the individual's capacity to acquire and use the information to adapt to environmental demands [11]. However, cognition can be challenged, and cognitive impairments can surface [12]. Cognitive impairments may be associated with cognitive decline due to aging, more-serious decline as dementia, or can be the consequence of an ABI. The latest is the category of patients involved in this study.

ABI is brain damage acquired after birth. The causes of ABI can be from non-traumatic brain injury (i.e., stroke, brain tumors) and traumatic brain injury (i.e., accidents, falls). It can affect cognitive, physical, emotional, social, or independent functioning. The consequences can vary from mild to severe [13]. We specifically focus on those people who have been affected in mildly their cognitive functioning. Eghdam, Scholl, Bartfai, and Koch [7] and Nilsson, Bartfai, and Löfgren [14] use the term Mild Acquired Cognitive Impairment(s) (MACI) "*to describe a subgroup of patients with mild cognitive impairment(s) (MCI) who are expected to reach a stable cognitive level over time. This patient group is generally young and have acquired MCI from a head injury or mild stroke*" (ibid). As this is similar to the user group we are interested in this paper, we borrow their terminology.

Cognitive impairments often persist after the ABI, and they can significantly affect an individual's abilities to perform everyday tasks, fulfill previous roles, and maintain personal-social relationships [14][15][16][17]. Thus, people's life becomes challenging, based on the severity of their injury. These people can experience difficulties in cognitive and emotional processing while having no or limited movement disorders and being independent in self-care [14][18]. Cognitive impairments after a brain injury can cause the following problems to people:

- They can feel a lack of energy for doing many things within a specific time frame.
- They can face problems with memory, which can be related to working memory or long-term memory. This creates difficulties in remembering and learning new things.
- They can face problems to keep the attention and concentrate on a specific issue and be able to eliminate the other unimportant information around them.
- They can face a reduced tempo of performing activities and engaging only those activities that are relevant to them.
- They can face a reduced multitasking capacity. The person becomes very tired if s/he has to manage too many activities at the same time.
- They can face an increased response time due to a slower process of thinking.
- They can face difficulties to start or initiate something, sometimes caused but a lack of motivation.
- They can face challenges in making mental changes from one topic to another.
- They can face difficulties in stopping, which relates to impulsivity, overactivity, and difficulties in controlling unwanted or inappropriate responses.
- They can face a reduced understanding of oneself, which involves a reduced ability to perceive one's own mistakes, to take into account the impression one makes on others, or to assess a social situation accurately.
- They can face thinking specifically and taking everything literally. This can also be associated with a loss of the ability to plan, look ahead, and think purposefully.
- They can have difficulties with using language and communicating, such as not finding a word, repeating the self in a conversation, having difficulties in making a point in a conversation, mixing words, or having difficulties in understanding humor or irony.
- They can face difficulties in processing the information received, solving problems, and executive functioning.
- They can face changes in how they engage in social life and social communication.

(the list presented above is a translated summary from [19][20])

The list of cognitive challenges that can be faced by patients that have had an ABI and suffer cognitive impairments shows how heterogeneous this user group is. In MACI people, these symptoms are mild, and the person in most cases continues having an active life, working or participating in social activities, running a house, or following hobbies. However, their life is not the same [21].

The symptoms of MACI patients mentioned above have similarities with other user groups as dementia and amnesia when memory problems are present, aphasia when communication problems are present, executive functioning, memory, attention, reduced tempo, and slow reactions are elements that are also associated with getting old. Moreover, problems with problem-solving can be similar to other cognitive impairment that patients have had from birth.

In the following subsection, we will describe the case of cognitive rehabilitation as the setting on which we conducted our research. In this paper, we have used interchangeably different variations to refer to people living with MACIs, such as "people with MACIs, MACI people or MACI person". Instead, from now on, we will use "patient(s) with MACIs, or MACI patient(s)" to refer to people with MACIs who are patients in a rehabilitation institution, which is the setting of our study. Moreover, in some cases, we will use the term patient(s) even when referring to the home context because people with MACI have a chronic disease, which makes them chronic patients in a rehabilitation context. When we describe our case, present the findings, and discuss them, we use the term "patients" as it better represents our participants. Instead, when we present the final guidelines, we refer to MACI people in general in any setting where they can be involved in designing new technologies useful to them.

B. *MACI Treatment – Cognitive Rehabilitation*

"*Cognitive rehabilitation can be defined as a learning experience aimed at either restoring impaired higher cerebral functioning or improving performance in the real-world using substitution or compensation techniques*" [16]. Cognitive rehabilitation is offered in specialized rehabilitation institutions. The cases presented in this paper are related to two projects that we did in collaboration with the Department of Cognitive Rehabilitation (DCR) of Sunnaas rehabilitation hospital in Norway.

Sunnaas offers multidisciplinary rehabilitation to people with complex functional impairments following illness or injury. We focused only on the cognitive rehabilitation process. The rehabilitation at the hospital (inpatient) is carried forward by a multidisciplinary team that helps the patient to define realistic and attainable rehabilitation goals and then define, in collaboration with the patient, a set of interventions that the patient should do for achieving the goals. The goals and respective interventions make a rehabilitation plan. The rehabilitation plan is imprinted in the "Goal-Plan" document. This Goal-plan is at the core of the rehabilitation process in the hospital. It helps to coordinate the activities that both the patient and the multidisciplinary team get involved in during the patient's hospitalization period. In every activity at the hospital, both the multidisciplinary team and the patient should refer to the Goal-Plan. For the team, the Goal-Plan resides in the hospital Electronic Medical Record system. Instead, for the patient, the Goal-Plan is a printed paper document. The patient continues with the rehabilitation plan at home and returns to the hospital after 2-6 months for short follow-up and further adjustments of rehabilitation goals. A more detailed description has been presented in Becker, Kirmess, Tornås, and Løvstad [22].

As part of the rehabilitation plan, different kinds of rehabilitation therapies, customized to each patient's needs, are applied. Some suggestions of rehabilitation techniques for working with patients that have had an ABI retrieved from the literature are:

- When communicating with an MACI patient, it is important to give her/him time, use short and concrete sentences, be clear, and try to receive confirmation that s/he understood what was said. Moreover, it is relevant to pay attention to how the patient feels that day.
- In order to increase understanding, remembrance, and better communication with MACI patients, repetition is necessary.
- Giving patients more time to do things or discuss is another strategy suggested. Moreover, considering the low capacity and high level of fatigue, pauses are recommended [19][20].
- In order to improve executive functioning Haskins, Cicerone, and Trexler [23] highlight the Goal-Plan-Do-Review (GPDR) model as helpful. GPDR requires the patients to increase awareness toward a specific goal that they want to achieve, plan carefully on how to achieve this goal, do the activity, and evaluate that after.
- In order to improve memory deficiencies, Haskins, Cicerone, and Trexler [23] describe different techniques which are in analogy with a "divide and conquer" approach where a bigger task is divided into smaller steps, and these steps are used to train by repetition. Considering that memory is fragile, an errorless approach in the rehabilitation techniques is used [18]. This aims to teach the patients only the right things so s/he might not risk preserving the error instead of the right information. Visual cues are also considered useful in helping to train memory deficits [24]. Cicerone et al. [25] suggest using references from patients' daily life because it is easier and more likely to remember the information when it is silent and personally meaningful.

The techniques mentioned above were both presented as in individual therapy sessions or group therapies. Patients with MACIs can find group therapies specifically helpful when it comes to discussing their problems and expressing more about their story in front of other patients with similar challenges. Now that an overview of the patient group symptoms and the rehabilitation process which s/he goes through has been presented, we further describe PD and how that might be relevant for this user group.

III. PARTICIPATORY DESIGN

PD was established at the end of the 1970s to democratize both the working life and the design process of new information technologies [26]. PD emphasizes the idea that those who will be affected by the design of new information technologies or digital artifacts, should get involved and have a say during the design process of these technologies [27]. PD considers users "domain experts" of the realities in which they live, so they must undertake the role of the designers [27].

In Routledge Handbook of Participatory Design, Simonsen, and Robertson [8] define PD as:

"*a process of investigating, understanding, reflecting upon, establishing, developing and supporting mutual learning participants in collective "reflection-in-action". The participants typically undertake the two principal roles of users and designers where the designer strives to learn the realities of users' situation while the users strive to articulate their desired aims and learn appropriate technological means to obtain them."*

At the core of PD is the idea of genuine participation in decision making. Genuine participation stands on a political rationale where the voice of marginalized groups is heard in the decision making that will influence them. Thus, designing technologies for patients with MACIs require their participation in the design process. Their marginalized voices in a paternalistic healthcare system where the patient follows what the doctors say should be raised and heard. By applying PD, these patients can have a say and genuinely participate in the design of new technologies, which will be used by them.

Moreover, as we describe above, MACI is a silent condition and often overlooked. This makes this group marginalized for the technologies that have paid more attention to more severe cases. We use the general term "technology(s)" because, in every technology type, digital solutions, tangible solution, ICTs (Information and Communication Technologies), etc. people that will be influenced by the new technology should be involved in its design. Moreover, different technologies can help MACI patients in different situations. For example, a tangible alarm button can be used in the case of a fall. Instead, an application can be designed to serve as a calendar.

PD is applied as a set of general principles that should be adapted to the specifics of the project. One of the principles of PD is "equalizing power relations". Thus, PD is concerned with questions such as "whom we may risk leaving out of the design space, how we can act upon such challenges and how to provide for alternative perspectives on participation and democratization" [28], and finding ways to give voice to those who may be invisible or weaker in organizational power structures [29] and beyond, thus building a democratic process.

Greenbaum and Kensing [29] point out that democracy is often propagated as a concept that is assumed to happen by itself. However, it requires educated and engaged people acting on their own interests and in the interests of the common good. Another principle of PD is mutual learning, where both designers and users should learn from each other in the design process. Learning also is cognitively demanding. Thus, the requirement for active participation and mutual learning poses a requirement for people to have a certain level of cognition. The number of people who live with cognitive impairments is high. To have a truly democratic process, we should not risk leaving this user group out of the design process.

Greenbaum and Kensing [29, pp. 33-34] have listed tools and techniques among the principles in PD. Different tools and techniques are used by designers and researchers to involve users in the design process. The designer should reflect which tools and techniques fit her/his user group and then adapt them with the practice at hand [30]. Choosing the right techniques and tools is as relevant as challenging when the focus group of the research is people who suffer cognitive impairments. A technique's goals, structure, and participation model can pose cognitive requirements to the participants. Some aspects of these requirements are fundamental to the activity (for example, a group discussion must involve communication), while other aspects are flexible (communication can be verbal, gestural, etc.). This flexibility can suggest ways in which a technique may be modified or adapted to people's cognitive abilities. However, what tools and techniques are chosen in a PD project would influence the true level of participation of the users in the design process [29].

A. Participatory Design with people with ABI

In PD, the active participation of people with disabilities in designing new technologies has been discussed significantly in conferences and workshops. Dementia patients or specific severe clinical conditions affecting cognition have been in the focus of many publications (examples [31][32][33][34]). Regarding the mild cognitive impairments, extensive research has been done with old adults or people with intellectual disabilities. However, little research has been conducted specifically on patients suffering from MACI [1] [35].

Augstein, Neumayr, Ruckser-Scherb, and Dielacher [36] have designed an interactive tabletop in the rehabilitation setting of people after an ABI by using a PD approach. However, they have involved in the design rehabilitation specialist as proxies of people that have had an ABI.

More research has been conducted with people that had a stroke. Balaam et al. [37] and Threatt et al. [38] have described some cases of involvement of stroke patients in the design process. However, this involvement is mostly in terms of interviews and observations and further during testing. The design is handled either by the designers or designers and clinicians together. The involvement of clinicians in design is as well described in Faria and Sergi Bermúdez [39]. Instead, Magnusson et al. [40] describe a case where stroke patients are involved in focus groups and as well in co-design workshops were brainstorming, and body-storming techniques are used.

Table 1. List of guidelines for conducting PD with people with cognitive impairments

No.	Guidelines	Reference
DG_PP1	Get the consent of the participant on various moments throughout the research process	[9]
DG_PP2	Communicate about project goals without intermediaries	[9]
DG_PP3	Give yourself enough time for general practicalities	[9]
DG_PP4	Get to know your target group, try to understand their cognitive deficit and become sensitive to their needs and situation	[9][41][42]
DG_PP5	Assess abilities through standardized tests	[4][9][43]
DG_PP6	Recruit and plan well in advance	[44]
DG_T1	Involve users in design in appropriate and familiar environments • The location should hold an appropriate social status • The choice of location should take into account the deficits of the participants and ensure easy access to the meeting room • Using the person's home might help to make the participant feel at ease	[9][45][46] [47]
DG_T2	Adapt the language to participants	[2][9][41][44] [48][49]
DG_T3	As the verbal might be a problem, make use of non-verbal elements such as visual stimuli like photos of objects or physical artifacts (notes, etc.)	[9][50][51] [52][53][54]
DG_T4	Use distinctive contextual cues (like nametags)	[4]
DG_T5	Consider the fidelity of the tools used to design • Consider using physical artifacts	[4]
DG_M1	Consider activities that facilitate challenges in envisioning future solutions • Boost sharing personal experiences [9][41][44][46][55][56] • Support the building of IT literacy [57] • Provide more hands-on activities and collective prototyping [55][58][59] • Use visual cues [41][42][55][59] • Make use of fictional characters in different scenarios [60] • Try to avoid appealing to the person fantasy; avoid too much choice [9]	[2][9][42][44] [48][54]
DG_M2	Consider activities that facilitate challenges with abstract concepts • Explain technological concepts in an easy-to-understand way [41] • Consider the fear of sketching [61] abstract ideas • Let the designer do the sketches if needed as a start for discussion [62]	[41][50] [59]
DG_M3	Involve people in designing valuable solutions, real purpose, interesting	[9][44][63]
DG_M4	Plan activities to surpass challenges in continuity • Using a document design history or summary documents of each session	[41][43][51] [64]
DG_M5	Use previously known activities and natural tasks	[9][55][59]
DG_M6	For each activity create a relatively open artifact and brief	[48]
DG_M7	Provide alternative activities so to engage all participants • Adapt methods so that it will take into account the difficulties in the comprehension and production of language, both verbal and textual • Adapt methods so that it can overcome impairments of memory • Facilitate participants to stay on track based on their individual needs	[9][48][50] [65][66]

No.	Guidelines	Reference
	• If working in a group, modify the method considering the different impairments each member of the group is facing • Consider the possibility of organizing individual participatory design sessions • Consider personalization and individuality for each patient	
DG_M8	Consider activities that can be flexible and empathic to adapt to the needs of the group: • Activities that can help create a friendly environment [44][46][67] • Activities that can boost participants self-esteem and confidence [52][68] • Activities that can include elements of playfulness [42][52][55]	[9][42][49] [50][54]
DG_F1	Researchers should clearly explain the purpose of events and the role of the participants	[9]
DG_F2	To enclose personal info from the facilitators will boost participants confidence and make it easier to share things	[9][69]
DG_F3	Incorporate Structure and Review in activities • It is important to foresee enough time for participants to get to know each other, repetition and constant reviewing of the different research/design phases	[2][9][43][46] [54][59]
DG_F4	During a participatory design session, try to minimize distraction and keep participants on focus	[9]
DG_P1	Consider a one to one or group work in a PD session • Try to overcome the challenge of working in groups [70] • Try to diminish the risk of the designer strong position [45] • Try to overcome deficits by pairing persons with different deficits into one subgroup [9]	[9][54][71]
DG_P2	Carefully decide the session duration and number of participants • In people with dementia [9] suggests working in small groups	[2][9][46][52]
DG_P3	Involve caregivers as support in conversation with participants • Third-party involvement [56][59]	[2][9][41][51] [59]
DG_P4	Involve caregivers as domain experts in the design process	[41][51] [72]
DG_P5	Eliminate usability problems with the carers of the patients • Specialist are relevant in generalizing solution [51] • Use persons who do not suffer from a deficit to get rid of general design problems [2][9]	[4][9]
DG_P6	Promote the involvement of family members; However, the involvement should be associated with a critical attitude	[73][74]
DG_A1	Try not to over-analyze the utterances of the participants	[9]
DG_A2	Be critical towards the representativeness of the participants	[9]

B. Participatory design with people with cognitive impairments

Research in PD with people with cognitive impairments is not novel. Researchers have been conducting PD with people with dementia, aphasia, amnesia, stroke patients, or in general, in older adults' populations, etc. for more than two decades. A set of workshops focused on people with cognitive impairment has been taking place in conferences [33][73][74][75]. Moreover, a considerable number of papers had been published [4][9][33][44][76]. The papers are mostly focused on describing and analyzing single PD projects and deriving implications for involving people with cognitive impairments in the participatory design process. Hendriks, Truyen, and Duval [9] present a broad review of conducting PD with people with Dementia and list a set of guidelines for participatory design together with persons with dementia. In their review, they have looked into PD projects with other user groups that have similar symptoms as persons with dementia such as amnesia, aphasia, and elderly. We build on their findings and expand them further with some findings from a literature review that we conducted in January-June 2019. The findings from the literature review are presented in Table 1.

As stated above, as a research question in this paper, we investigate what a researcher should take into consideration when conducting PD with people with MACIs. However, as mentioned above, the number of publications specifically related to this user group is low, and only a few other publications refer to PD with stroke patients. Hence, in order to have a broader overview of how to conduct a PD project with our user group, we did a systematic search in the literature of conducting PD with people with any kind of cognition challenges considering the similarity in symptoms.

We limited our search to the ACM and Springer databases since they are the main publishing venues for conferences and journals in design. We initially planned to investigate publications in specific conferences. However, in order to mitigate the risk of leaving out any relevant publication, we decided to expand the search generally for the two databases. We used several search terms in both databases. The constant search term was "participatory design" or "codesign" (and variations of it), qualified by more specific searches for user groups that experience cognitive impairments such as older adults, people with dementia, aphasia, cognitive decline, brain injury, and stroke.

Moreover, we added search terms that relate to the symptoms that people with cognitive impairments face, such as memory, attention, and tiredness. The search gave us an extensive set of papers. After removing duplicates, we ended up with 326 papers in ACM and 146 in Springer. The initial phase was to read through the abstract and quickly scan the papers' headings if there was any part that was dedicated to PD or the design process. Based on this, we decided if we needed to read the paper further. From the first scan, we decided to read 105 ACM papers and 38 Springer papers thoroughly.

After this, we discarded papers found to be out of scope. Some because they did not explain a specific case of participatory design with people facing some form of cognitive impairments, some for just mentioning a PD process without additional information, and somewhere the design process was not a true participatory design process with the involvement of users as partners in design. We concluded the selection with 105 papers. We classified the papers into Technology papers (18); Methods papers (49); PD practice papers (28) - in which the PD process and the tool developed were explained; and General PD challenges papers (17) - focused more in a conceptual discussion of what implications and what concepts are important on co-designing with people with disabilities. It was not a precise positioning of the papers in one of the categories mentioned above, and some could belong to all. However, each paper was assigned to only one category based on the paper's main contribution.

To analyze the data from the literature review, we were guided by content analysis and grounded theory approaches [77]. The content analysis starts by assigning specific descriptors to blocks of text in the collected data, a process called "coding". The coding can be emergent or a priori. The emergent coding is the core of the grounded theory approach, and the codes emerge from the data under review. Apriori coding involves the use of an established theory or hypothesis to guide the selection of codes. These categories might come from previously published work in related areas, or own prior investigations of the topic at hand.

We started by openly coding our data. Codes that described similar guidelines were grouped into concepts. Concepts were further grouped into categories (axial coding). Throughout the coding process, both authors of the paper were engaged in intensive discussions to find meaningful codes or concepts (i.e., topics).

To compare the results and refine and enrich our emergent codes, we used apriori coding in the second round of analysis. The apriori codes used derived from Hendriks, Truyen, and Duval [9]. Moreover, Hendriks, Truyen, and Duval [9] have categorized their guidelines. We preserve these categories and introduce a set of guidelines and actions to take for each guideline suggested by different authors (cross-checked with the guidelines from Hendriks, Truyen, and Duval [9]) for each of the categories.

In Table 1, we have listed guidelines for conducting PD with people with cognitive impairments, and group these guidelines into the following categories: Preparation (coded as DG_PP), Tools (coded as DG_T), Moderator (coded as DG_F where F stands for the facilitator, inspired by [78]), Methods (coded as DG_M – instead of methods we will refer to this category as Techniques inspired from the Simonsen, and Robertson [8]. We save the letter M in coding as the letter T has been used for the Tools category already) and Participants (coded as DG_P) and Analysis (coded as DG_A).

IV. METHODS

In this section, we initially present the data collection method, followed by the method used for analysis.

A. Data collection

The study aims to define a set of guidelines for conducting PD together with people with MACI. For defining the guidelines, we build on two empirical cases where we conducted PD with people with MACI.

In the first case, the aim was to redesign the Goal-Plan document used as a central document of patient rehabilitation at the DCR. The redesign aimed to make the document more useful for the patients during their rehabilitation.

Table 2. A summary of activities on how empirical data was collected

	Workshops	Participants	Recordings	Facilitators	Reflection-on-action	Preparation meetings
Workshop Outline 1	Workshop 1	4 patients	1h:10m	2	Facilitators + clinical representatives	3
Workshop Outline 2	Workshop 2 Workshop3	4 patients 2 patients	1h:10m	2 2	Facilitators + clinical representatives	2
Workshop Outline 3	Workshop 4 Workshop 5	2 patients + 3 staff 5 patients + 5 staff	1h:45m	2 2	Facilitators	2 2

We participated in the project as researchers pursuing our aim of investigating how to involve MACI patients in PD and designers to help the DCR to redesign the Goal-Plan.

Different methods and techniques can be applied in PD projects to involve the users in design. In the PD handbook, Brandt, Binder, and Sanders [30] describe a set of techniques that emphasize different parts such as telling, making, and enacting or the possible combination of these. We decided to involve MACI patients in PD workshops as our method of investigation. Workshops give the possibility to evaluate different traits of user involvement in the design, thus giving us the possibility to explore more on MACI patients and how to conduct PD with them. In a workshop, several techniques can be combined.

For redesigning the Goal-Plan project, we worked together with a project committee with representatives from the multidisciplinary team at the DCR. We prepared and conducted three workshops with a total of 10 participants. All three workshops were audio recorded. Moreover, after the workshops, a reflection meeting was conducted between facilitators and clinical representatives to discuss the workshop and the participants' engagement. The first author kept notes from these meetings electronically.

The second case is a continuation of the first case. The aim is to design a digital Goal-Plan that can be used by patients and staff in the process of defining rehabilitation goals and keeping track of the goals throughout the time that the patient is at the DCR. Thus, the digital tool would work as an enhancement or substitution of the current paper document. Two workshops with people with MACI, where a total of 7 patients and 7 multidisciplinary team members participated, were organized as part of the project.

The authors of this paper were involved in the project as researchers and designers. Both authors worked in the preparation of the workshops. The first author was as well a facilitator in the workshops. Both workshops for this digitalization of the Goal-Plan project were audio-recorded, and the reflections from the workshops were as well audio recorded. A summary of the data collection is presented in Table 2.

B. Analysis

To analyze the experience of conducting PD with people with MACI, we build on Schön's [79] approach of the reflective practitioner. Schön describes two types of reflections:

Reflection-in-action is undertaken in the indeterminate zones of practice. The reflective practitioner "*thinks up and tries out new actions intended to explore the newly observed phenomena, test tentative understandings of them, or affirm moves invented to change things for the better. What distinguishes reflection-in-action from other kinds of reflection is its immediate significance for action*." ([79, pp. 28-29]). Referred to as a reflective conversation with the situation.

Schön's use of the term reflection-on-action refers to the process of making sense of an action after it has occurred. It serves to extend one's knowledge base.

The two PD projects described in Sections V and VI have involved both reflections in action and reflection-on-action. Some reflections for the first case (Section V) have already been presented in a previous publication [1]. Those reflections were made before the literature review. The reflections presented in this paper have been refined further due to the findings in the literature review.

The analysis was conducted in two parts. Initially, for each of the cases, the authors, based on the experience of conducting PD with MACI patients, made a list of reflections on each of the workshops. The list of reflections for the three workshops in Case 1 started in June 2018 and was published in March 2019 [1]. However, we expanded the list of those reflections in January-February 2020, where additional elements of the workshops or the preparation phase were listed as reflections. For Case 2 and its two workshops, the list of reflections was made in February 2020. The reflections list is what Schön [79] defines as reflections-on action. The reflections on action presented in this paper are the ones from the perspective of the facilitator in the PD sessions and not through an evaluation of the sessions from the patients' perspective.

We present the reflections for each workshop outline because the same outline had mostly the same reflections. If there were different reflections among workshops in the same outline, we have made sure to capture and include it in the reflections by referring specifically to the workshop.

The second part of the analysis was the refining of the reflections drawn on existing literature guidelines for conducting PD with people with cognitive impairments. The initial reflections list for each workshop outline was put in an excel sheet where each reflection was inserted in a different column. Instead, in the rows, we listed the guidelines found in the literature (as in Table 1). Figure 1, the under excel sheet, shows this part of the analysis. We started a qualitative "correlation analysis" based on reflections and interpretations. We initially went through the guidelines from the literature and analyzed which of our reflections from the empirical cases were compatible with specific guidelines found in the literature. Some of the guidelines in Table 1. helped the authors reflect on topics we had taken for granted in the initial reflective analysis, leading to updates in our reflections' lists.

Moreover, a second round of reflective interpretative qualitative "correlation analysis" was undertaken. Again, we read through the guidelines from the literature and analyzed which of the reflections were compatible with them, leading to further refinements of the reflections list.

In the third round, we started grouping our list of reflections based on the categories that we had in the guidelines from the literature in Table 1. These categories are borrowed from Hendriks, Truyen, and Duval [9], and are Preparation, Tools, Techniques, Facilitators, Participants, Analysis. We present the reflections for each workshop outline divided into each of these categories.

In the fourth round of analysis, we created a new version of the excel file and created separate sheets for each of the categories, as shown in Figure 1, the top excel sheet. The reflective interpretative analysis in this phase led to a set of guidelines for conducting PD with MACI patients, emerging from the literature of PD guidelines with people with cognitive impairments and the experience of conducting PD with MACI patients in two projects.

The correlation analysis was also influenced by the characteristics of cognition challenges that the MACI patients have and the suggestions coming from rehabilitation theories, as described in the "Cognitive Rehabilitation Manual" [23]. The final list of guidelines for conducting PD with MACI patients is included in this paper's discussion.

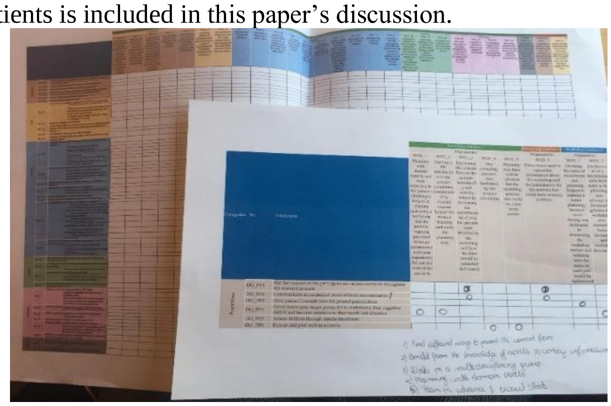


Figure 1. Visualization of the excel sheet used for the analysis and the smaller sheet of analysis for tools guidelines for conducting PD with MACI patients

V. CASE 1: REDESIGNING THE GOAL-PLAN

Above we presented the cognitive rehabilitation process in a hospital in Norway. The structure within the hospital, which is specialized in cognitive rehabilitation for patients with MACI, is the DCR. One of the primary working documents at the DCR, as explained above, is the Goal-Plan. Intending to empower the patient, the DCR wanted to redesign the layout of the document so it would fit more patients' needs and consequently make the patients make more and better use of the document during their stay at the hospital. We will refer to this project in this paper as "the redesign project".

As stated above, the authors were involved in the project in the role of researchers and designers to investigate patients' needs and, together with the patients, redesign a new version of the Goal-Plan that would fit those needs. Both authors worked in the preparation phase and the reflective analysis presented in this paper, and the first author participated and facilitated the workshops described below.

In collaboration with a project committee with representatives from the multidisciplinary team at the DCR, we prepared and conducted three workshops with the patients. The title of the workshops was: Redesign the Goal-Plan: A patient's perspective. The workshops aimed to get an understanding of what experience the patient has had with the Goal-Plan and discuss ideas on how to redesign that document so that patients can integrate it more in the activities during their rehabilitation period at the hospital.

In total, ten patients participated. In the first two workshops, patients that had been at the hospital for more than a week were invited, so they were familiar with the document to be redesigned. Four patients participated in the first two workshops. In the last workshop, only two patients participated, who were back at the hospital for their follow-up week, six months after their discharge. We were two facilitators – the first author in the role of facilitator and designer and as well one representative from the hospital who had an experience of working with this patient group but was not their direct therapist at the DCR. We refer to this second facilitator as a "knowledgeable third party" [1].

Below we describe the preparation process for the workshops and present and reflect on two workshop outlines

A. *Preparation*

In the preparation phase, we collaborated with a multidisciplinary team at the DCR. This team was assigned as a leading committee for the project, and we refer to it in another publication as the multidisciplinary project committee.

The committee was compounded by therapists who had high expertise in working with MACI patients. We refer to these people as "domain experts".

In order to facilitate the collaboration in the planning phase, prior to workshop 1, the first author conducted a PD workshop with the multidisciplinary project committee. The aim was to introduce the committee to some of the techniques used in PD and also for the first author in the role of researcher and designer to learn more about the user group. The workshop was initially seen with skepticism from the committee members, but once they understood the aim, they expressed that it was useful to learn about the PD approach and be able to contribute to *"design"* better workshops to enhance MACI patients' engagement.

Based on an agreement with the project committee and some shared editing iteration, the DCR staff members gave patients before the workshop a consent form, including an invitation to the workshop and a description of the project. The consent form was written in a very simple language to make it easier for the patient to follow. However, it was a detailed and consequently long description, to make sure that all the ethical issues were covered.

The project committee members tried to recruit their patients, asking if they would be interested in participating in the project. The patients that showed interests were then presented with the project invitation and the consent form. We were aware that the description might be excessive for the patient and could make her/him neglect reading it carefully. To make sure that the patient understood the consent form, one of the staff members at DCR spent time with the patient (that had expressed the willingness to participate) before the workshop, going through the document, and provide further explanations where needed. The workshop time was included in the participant's weekly plan. This is a calendar document where all the activities of a patient at the hospital are listed. The participants said that listing the workshop in the calendar helped them remember.

Regarding ethical concerns, we decided together with the committee that no personal patient data would be recorded. However, the sessions would be audio recorded so we could analyze the data later. The data collected through recordings are considered not anonymous (they are identifiable data), so they would be stored carefully in a safe location. In accordance with the project agreement, all digital data was stored in a personalized folder at the hospital servers.

During the first workshop, we noticed that some patients did not have a clear idea of what they had committed and needed a reminder. Thus, before workshops 2 and 3 in the morning of the workshop day, the nurses at the DCR talked with the patients again, to make sure the patients understood the scope of the project and were reminded and more prepared for what they were participating in later in the day. This was especially useful for patients with memory problems.

Moreover, we decided to conduct the workshops in environments that were familiar for the patients at the unit and would be adequate not to distract the patient and influence in their attention.

B. *Workshop Outline 1*

We organized Workshop 1 in two parts.

The first part was "storytelling". The title was "Sharing your experience". We invited participants to talk about their experience with the Goal-Plan. Moreover, we provided participants a whiteboard with a print out of the old Goal-Plan was set in the middle and sticky notes in different colors. The patients could use those to write down keywords to facilitate remembering what they had to say when their turn would come.

In the second part, we challenged participants to think, "How would they want their Goal-Plan to look like" and design the idea afterward. The technique chosen was drawing and discussion.

For the second part, we removed the Goal-Plan and gave each of the participants a white sheet of paper, where they could design their ideal Goal-Plan.

The first author led the session. The second facilitator was facilitating the communication with the patient when that was needed.

C. *Reflections-on-action*

In this subsection, we present a set of reflections-on-action from workshop outline 1. Reflections are grouped based on the categories borrowed from Hendriks, Truyen, and Duval [9] and also used in Table 1 above. The reflections are presented as insights on best practices or problems we faced based on a look back on our experience while conducting PD with MACI patients. We refer to these reflections again in the discussion section when we analyze these reflections from a theoretical perspective and conclude with a set of guidelines for conducting PD with MACI patients. Numbering the reflections has no specific meaning, but it useful to refer to after in the discussion.

Preparation

- WO1_1 – Involving domain experts in planning the PD process for the MACI patients resulted in arranging a workshop which was enjoyable for the participants and provided the researchers with relevant information. Moreover, for a researcher or designer would be very difficult to gain deep knowledge of the patients' needs in the timeframe of a project. Thus, having in the planning team, the domain experts facilitates having a broader perspective for the patient group and planning better.
- WO1_2 – Having a PD introduction to the project committee compounded by domain experts helped the mutual learning and made the planning easy. We, as designers and researchers, learned more about the user group, their needs, and challenges. During the mutual learning period, the first author had in situ conversations with the domain experts about the patients and how to design with MACI patients. Relevant tips and literature were exchanged in the conversation. Moreover, the domain experts gained new knowledge about the PD techniques and gave more specific recommendations on the activities that could be adequate or not for the workshop participants.
- WO1_3 – Explaining the consent form individually and verbally to the patients that had shown interest to participate in the workshop helped in increasing their understanding of the workshop's aim and how the data would be collected and stored. The lack of a verbal explanation would have caused one of the participants not to understand what was signing for. The person was diagnosed with dyslexia.
- WO1_4 – The project committee facilitated the recruiting process. The patients were already in the hospital unit, so it was easier to invite them to participate. However, in order to establish a relationship with the hospital and receive permission to conduct research in the DCR, it was a long process. In this process, we had two significant gatekeepers that supported us.
- WO1_5 – Planning was done well in advance, but the recruiting process was mostly done on a one week before notice. This because the persons invited were at the hospital, which was easily accessible for the authors. However, the process was limited in time because the persons would be patients at the hospital only for five weeks. Moreover, the domain experts suggested that even if recruited well in advance there is a risk that the patient might forget about the workshop date or in another scenario, the change in his/her condition from one week to another can influence them to drop the participation (some can feel too tired, not motivated, etc.).

Tools

- WO1_6 – Involving the DCR patients in a workshop at the hospital, which was a familiar environment for the patients, created an easy-going atmosphere. We used a room the patients had been in before for some of the activities at the hospital. Moreover, the room was easily accessible to everyone.
- WO1_7 – We distributed the materials on the table. The participants liked to look through these carelessly distributed workshop materials on the table, but they did not use any of them. Any of the four participants did not embrace the unstructured and open way of presenting materials.

Techniques

- WO1_8 – The participants did not use the sticky notes at all in the first part. They instead expressed themselves verbally and shared their stories easily with the others—three of the participants engaged in talking about their stories and their Goal-Plan in the first part. Instead, writing down in sticky notes, the main points in their story was difficult, and the participants seemed uncomfortable to do. Thus, the facilitators dropped the idea and instead investigated more through in situ questions. One of the participants was more reluctant to share his stories. After the workshop, we found that the participant's cognitive impairments had influenced his ability to articulate himself.

- WO1_9 – The fear of white paper, the blank page syndrome [31][32], was made visible in the second half of the workshop. The participants were good at articulating their needs verbally, but they were not able to create a visual image of their needs and consequently design ideas. Participants received the white paper and felt uncomfortable to write something on it. Instead, they started telling the facilitators what they wanted to have in a new version of the Goal-Plan. Both facilitators, ones noticing the hesitation, immediately reacted and suggested the participants to not focus on the white paper but more telling us their ideas about a new Goal-Plan. All the participants received the white paper with them, and only one of the participants came back the next day with a design suggestion and talked personally to the first author. Joyce [33], in her dissertation, discusses the role of open options in creativity and finds how the openness of the design space can constrain creativity. That is what we experienced with the MACI patients.
- WO1_10 – Realizing the hesitation on writing in the sticky notes and designing on the white paper, both facilitators abandoned the writing and drawing idea and started bringing up the questions of the workshop as discussion points to elicit ideas and needs from the patients. This flexibility toward the participants' needs and comfortability made it possible to end the workshop with some interesting data from the participants and, most importantly, have an enjoyable session for the participants, which expressed the enjoyment to the facilitators.
- WO1_11 – The broad approach of the workshop was distracting for the participants. Opening the workshop by asking the participants to talk about their experience with the Goal-Plan resulted problematic. It made them focus more on their goals and their specific problems rather than the main project aim, the Goal-Plan layout. Noticing the distracting reaction that the initial more general questions about the Goal-Plan created, the facilitators started asking the participants more specifically about the Goal-Plan layout. This resulted in participants being more focused on the main aim of the workshop and giving feedback specifically for the issue at hand. In the planning phase, the domain experts warned facilitators about the patients' willingness to tell their story and the possibility of getting distracted and get lost in detail. We experienced this with our participants. The presence of the second facilitator contributed in improving the communication and bringing the participants in focus.
- WO1_12 – The participants had different MACI, which meant they had different cognitive challenges. This required that within the aim of the workshop to adapt to each of the participants' cognitive needs. We noticed that two of the participants were more focused on discussing personal goals than contributing to the layout of the document. Thus, to them, we started asking more specific questions with short, clear sentences. Instead, one participant was more reluctant to share his stories. Thus, we tried to ask indirect questions so the participant could start sharing some ideas, and also, we were careful not to push the participant out of the comfort zone.
- WO1_13 – The aim of the project concerned the participants directly. Thus, they were interested in contributing as much as possible. One of them expressed that the contribution to designing a new Goal-Plan would help maybe not them directly but others after them in their rehabilitation. The same participant had discussed the workshop in the evening with the kin, and they had together designed something in the white paper.
- WO1_14 – Building on the participants' personal experience and opinion resulted in being a positive experience because the participants where used at the hospital to share their stories with different therapists. Moreover, talking about the Goal-Plan from their experience created an environment that was friendly and boosted the participants' self-esteem as there was no right or wrong answer.
- WO1_15 – A thorough review of the literature about the patients' clinical condition as well as observing the patients in the unit, made clear that it is a very special user group. The committee suggested focusing on the patients' abilities and how to strengthen those abilities during the workshops. The staff highlighted the patients' willingness to share their stories and express themselves both through words and as visual imagery.
- WO1_16 – Based on the committee expertise, the optimal workshop duration would be 1 hour, divided into two parts, each of 20-30 minutes with a 5-10-minute break in between. In this way, it would be possible to have the patient concentrated all the time without fatiguing him/her. This was the case during the first workshop. The activities in each part lasted enough to not be overwhelming for the participants. The participants seemed happy after the workshop.
- WO1_17 – Creating a friendly environment with coffee and biscuits and long breaks was stated as a positive trait of the workshop from the participants. The had the possibility to talk more with each other and with the facilitators during the break. However, the participants knew each other from before as they had been involved in group therapies. This created a friendly relationship between them and, from our interpretation, influence positively during the workshop by boosting the participants' willingness to share their opinions and experiences.

Facilitators

- WO1_18 – Having as a facilitator, a "knowledgeable third party" improved the communication process for those participants whose communication ability was affected by brain injury. Discussing the issue of facilitators with members of the committee, we considered an extension of the workshop team by someone from the clinical side that knows how to work with the patient group but is not directly involved with the participating patients. The committee suggested a member from the Learning and Mastering Center at the hospital, which was specialized in providing patients with a deeper insight regarding their health. The member might have met the patients during other activities around the hospital but was not part of the DCR staff and not directly involved with the patients. The involvement of a person that fulfills this requirement as a facilitator in the workshop was very useful in smoothing the communication and boosting the participants' contribution.

- WO1_19 – Facilitator's challenges of not being fluent in the Norwegian language created a fun atmosphere that helped the participants to feel at ease and not be shy to highlight their challenges. Moreover, as a facilitator keeping up with a positive and humble attitude helped the facilitators to communicate with the participants better. Facilitators were speaking on a slow tempo and not in a feeling of rush. Moreover, facilitators used a simple language and showed respect for the participants' knowledge and experience with their situation.
- WO1_20 – The broad approach of the workshop diminished the direct participants' contribution to the design of the Goal-Plan document. Thus, the facilitators had to intervene to help the participants get back on track and focus them on the aim of the workshop.

Participants

- WO1_21 – The number of facilitators should balance the number of participants. We decided that two facilitators (the first author having design skills and the knowledgeable third-party having domain knowledge) would be sufficient in a workshop with four participants. This saved the balance during the workshop. The participants were in the majority, so they were not put in the spotlight, which could have created stress. However, having one leading facilitator helped to keep the focus in one direction.
- WO1_22 – The committee suggested that the maximum number of participants per workshop should be around 4 or 5. In this way, the participants would feel more comfortable and have the right space to share their stories and opinions. Indeed, that worked well in workshop 1 with the four participants. Each of the participants had a dedicated time to express her/himself.
- WO1_23 - Recruiting MACI patients at the hospital through domain experts made the process of recruitment easy. Moreover, the domain experts served as the gatekeepers for involving in the workshops MACI patients based on the standardized tests that the patient has had at the hospital.

Analysis

- WO1_24 – We conducted a reflective analysis with the two facilitators and representatives from the DCR. Analyzing the findings in an interdisciplinary group helped in making a better sense of the participants' behavior during the workshop (reflected in this paper) and their feedback (integrated into the new Goal-Plan design). The assessment of the participants at the hospital and knowing their diagnosis was useful in triangulating the findings and make out meaning from them.

After the first workshop with workshop outline 1, we reflected on the things that did not work perfectly during the workshop, and we made another plan for the next workshop. This leads us to workshop outline 2.

D. *Workshop Outline 2*

The workshop was organized as an updated version of a future workshop, as presented by Jungk and Müllert [80]. Future workshops have been widely used in PD. The aim is to make people critically discuss a current situation and then envision possible improvements for the issues critiqued in a fantasy phase. After a phase of envisioning any solution, it comes to the realization phase. In the realization phase, feasible solutions based on what the technology allows are discussed further. We ideated an updated version of the future workshop, as presented below.

The workshop was divided into three parts.

In the first part, the participants got a version of the old Goal-Plan printed out. Next, to each of the fields in the document, we added two icons, thumb up and down. We asked the participants to mark with thumb up those fields that they considered important for their rehabilitation. After choosing to thumb up or down, the participants were asked to share their choices with the others and tell a little bit why they decided so. The aim was that the participants could discuss the choices among each other and maybe build on the ideas of each other. To structure the discussion, the knowledgeable third-party facilitator started going from one field to another and asking participants for their choice. Thus, it made it easier for the participants to follow and contribute to the discussion.

In the second part, the participants were asked to try to rewrite the fields (words used in the document) that they found important, in a way that they thought would be easier to understand and read. The participants were suggested to add additional fields if they thought they would be important in the Goal-Plan. Again, they had to work initially on their own and then present their suggestions to others. The other participants could ask questions or comment. The facilitators asked questions to enable the participants to articulate more of their needs and sometimes ask questions to understand better what the participant meant, to not risk misinterpretation during analysis.

The third part was called "rearrange". In this part, the participants were asked to rearrange the fields of the document which they had selected and rewritten in the first two parts. They could arrange the document as they wanted, add new fields or, change the structure of the document. At this point, participants could use the template of the old Goal-Plan or get a white sheet and design on it individually. Colored sticky notes and pens were provided.

In the third part, the participants were also provided with some examples of designs made by the multidisciplinary project committee in the workshop with the designer. Participants could have a look at those sketches for a short period for inspiration. This exemplars' aim was to help participants surpass the white page syndrome.

The same workshop format was used in the third workshop, where two patients that came back for a follow-up week participated. The only change was that in order to customize the discussion for these two participants and adapt to their experience, we focused mostly during the workshop on how the continuity of the rehabilitation plan could be achieved when the patient leaves the hospital.

E. *Reflections-on-action*

In this subsection, we present a set of reflections-on-action from workshop outline 2. Reflections are grouped based on the categories borrowed from Hendriks, Truyen, and Duval [9] and also used in Table 1 above. Here we do not have any reflections in the category "Analysis", so this category is not included below. The reflections are presented as insights on best practices or problems we faced based on a look back on our experience while doing workshops 2 and 3 with MACI patients. Here we include only reflections that were additional in workshop outline 2. The reflections from workshop outline 1 had already been taken into consideration before planning workshop outline 2.

Preparation

- WO2_1 – In workshop 2, we found that repeating the information about the workshop and the participation for the patients that had committed to participate and were suffering from memory problems resulted in them being more focused during the workshop and more prepared on what they were going to discuss. The nurse, as in workshop 1, had a meeting with the prospective participants to help them understand the consent form. In addition, the nurse talked again with the patients in the morning before the workshop to remind them about the workshop. This brought a more engaged and wider feedback from the participants during the workshops.

Tools

- WO2_2 – Distributing the sticky notes and the colored pens in an organized way in each of the parts of the workshop was noted to stimulate the participants to engage with the tools. They did not have the hesitation to chose among the tools because they had a structured set of tools for each part of the workshop.
- WO2_3 – Having the Goal-Plan in a printed version in front of each of the participants individually as part of the workshop toolkit, helped them be more engaged with it. Participants could customize the printed piece of paper as they wished – they owned it. Moreover, using the workshop tools to collect data facilitated the interpretations. For example, we could count how many thumbs up or down were related to a field in the Goal-Plan.

Techniques

- WO2_4 – We implemented a task-oriented approach in the workshop. Each of the parts was framed as a clear task that would serve a specific purpose. Participants liked this approach. They engaged in significant discussions with each other and the facilitators. Moreover, they started building on the ideas of each other. If someone brought up a new idea that would also trigger a discussion among other participants.
- WO2_5 – We observed aiding the participants through specific cues such as marking thumb up and thumb down in the Goal-Plan, made it easier for them to starts the discussion and elicit their ideas. The usage of thumbs up and down was associated with more personal stories and individual opinions about why a specific field in the document was relevant or not. So, having a structured way where to start the discussion was useful in helping the patients to build up ideas and relate those to personal experiences. The cues included in the tasks facilitated participants' ideas and discussion. This created the opportunity for the facilitators to ask more questions to elucidate the meaning of participants' ideas.
- WO2_6 – Having a narrower scope of the PD session helped participants to stay focused and contribute significantly. While these reflections are not new, they appear very important in the case of patients with MACI.
- WO2_7 – In the critique phase of the future workshop, we did not organize a real critique session. Instated, we asked what the participants liked and what the participants thought needed to be improved further in the current Goal-Plan. Providing both the thumb up and down options enabled the patient to think that some things need to be improved, but at the same time, that there are other things that are extremely relevant and need to be preserved. This was useful to keep participants' good feelings and not expose them toward a negative mindset. Thus, considering ways of using positive rhetoric that can elucidate a critical perspective from the MACI patients in workshops is very relevant.
- WO2_8 – The second part of the future workshop is the fantasy phase. It was clear from the first workshop that the MACI patients could not produce much information while moving directly to the fantasy phase (hesitation of the white paper). Thus, before jumping to the fantasy phase, we introduced a transition phase by asking participants to rewrite some of the fields in the Goal-Plan that they thought could have been written better for them. Rewriting aided participants to start envisioning a better solution for the Goal-Plan but still be connected to the things that they knew, meaning to the Goal-Plan that they had seen many times. In another publication [1], we refer to this as "the teaser of future envisioning". The teaser is a simple known task that helped participants in transitioning toward the fantasy phase of the workshop and be able to design a new version of the Goal-Plan either by rearranging the old one or by designing in a white paper. The white paper syndrome was defeated, and participants could come up with design suggestions for a new version of the Goal-Plan. Figure 2 shows some of these design suggestions provided by patients.

- WO2_9 – The use of exemplars in the "rearrange" part of the workshop, might lead and influence participants' ideas. We were skeptical about the usage of these exemplars, but we wanted to observe what their influence could be and how the patients would react toward that. However, screening participants' designs did not reveal a noticeable influence from the exemplars presented. Based on our reflective analysis the reason the exemplars did not influence the design ideas was that they were presented to participants in the last part of the workshop, and participants had already built up a mental vision of their Goal-Plan in the previous phases. Moreover, we exposed the exemplars only for a short period and explained that they were supposed to be triggers for possible options of how a Goal-Plan could look like. Integrating exemplars was inspired by research through design and Gaver's work with the ludic design [81][82][83]. Looking at the amateur designs from the staff inspired MACI patients participating in the workshops to get the colored pens and sticky notes, and start designing, overcoming the fear of the white paper. However, this is a very delicate usage, and more investigation is needed.

Figure 2. Workshop 2 - Patient's design suggestion

Facilitators

- WO2_10 – Providing structure and review helped participants to give more ideas and articulate their thoughts more deeply. The knowledge third party going from one field of the Goal-Plan to another to ask if the participants had marked that with thumb up and down facilitated the initiation of the discussion among participants. Moreover, asking additional questions helped participants to express their ideas better.

Participants

- WO2_11 – In the third workshop, the two participants were of different natures. One of them was more expressive, and the other more reserved. Hence, the facilitator had to make sure that both were getting the same time and attention by providing the same time and attention to both participants.
- WO2_12 – Participants with different backgrounds but similar cognition levels seemed to work better with each other. This was noticed, especially in workshop 2. The four participants had different MACI but more or less similar functional level. This helped in keeping up at the same speed and feeling motivated by each other.
- WO2_13 – We noticed that participants were comfortable with the group work. They were used in working in groups, from previous group therapies at the hospital. Many showed during the workshop confidence point to another patient on things they thought were similar. Moreover, we noticed that some participants who had different ideas from the group felt confident enough to share them with others. Especially if the idea they had was related to a story in their life. However, the situation was different in workshop 3 when one of the participants was perceived as influencing the ideas of the other. Facilitators had to intervene through more strategic questions to retrieved more hidden ideas. A helpful thing was that each of the participants had to work initially on their own and then discuss with the others in the group in its own turn. This helped in preserving the individuality of opinions.
- WO2_14 – Participants that were present in workshops 2 and 3 had a milder ACI (acquired cognitive impairments). Thus, they could contribute better in giving feedback on design details as well.

VI. CASE 2: DESIGNING A DIGITAL GOAL-PLAN

In June 2019, as a joint collaboration among our research institution and the DCR of the rehabilitation hospital, we started "The interactive Goal-Plan" project. The project aims to develop a digital version of the Goal-Plan, which can support the patients to take more control over their rehabilitation at the hospital and outside it. We will refer to this project in this paper as "the digitalization project".

Figure 3. Participants working in pairs patient-therapist during workshop 5

The aim of supporting the patient to take more control over her/his rehabilitation starts with the patients deciding themselves what they would like to have in a technological tool designed for them. This philosophy of the hospital is compatible with the PD principles. Thus, a PD process started in January 2020, where the authors of this paper in the role of researcher and designers were involved in two PD workshops organized respectively 22nd of January and 5th of February 2020 at the rehabilitation hospital premises with patients with MACIs and staff from the DCR. These workshops will be described below.

A. Preparation

In this project, a multidisciplinary project committee was created again. The first author in the role of the designer and researcher worked in close collaboration with few representatives from the DCR. One member of the DCR, which was involved directly with the patients, was in charge of the recruitment and deciding on a venue and time suitable for all. The planning of the workshops was done in collaboration among the authors of this paper and an Occupational Therapist (OT) that had been working at the DCR before but was now in other duties at the hospital. This person had been a crucial person in the first case described above. Thus, she had created knowledge about the design process and the PD principles previously. As she was not directly involved with the patients but had high expertise in working with the patients, she represented the “knowledgeable third party” in this project. A more elaborate plan was made considering the experience of the facilitators and the lessons learned from the first case.

For workshop 4, the plan was to recruit 5 patients. However, only two patients responded positively to the invitation. The workshop was planned to have a room that could support groups working in pairs, but the room available had only one single long table. We needed to be flexible to the positioning of the participants, so not much noise was created and distract the patients. Workshop 5 was planned for the 29th of January, but we needed to postpone it due to a small number of patients willing to participate. Workshop 5 was organized in a room called the “Idea Lab” suitable for bigger group workshops.

Participants in this case workshops were patients and DCR staff members. Hence, we prepared two invitations and consent forms in order to make the call more personalized. The consent forms this time were shorter and more precise. It had a clear part highlighted in colors where the aim of the workshop and each of its parts were presented. This was followed by a description of how the data collected would be handled. The one-page invitation resulted in being easier for the patients to read through and understand. The “knowledgeable third party” mentioned above contributed substantially to writing the consent form. Moreover, together with the designer (first author), they drafted the workshop description and presentation. The presentation was planned to stay on the screen, and the written material was distributed to the participants as a cue for understanding better the tasks in each of the parts.

B. Workshop Outline 3

In workshop 4, two patients and three therapists from the DCR participated. The workshop was called “My Interactive Goal-Plan – Defining goals”, and the aim was to envision and discuss requirements for a digital tool that can support the patients to be more involved in setting their rehabilitation goals during their first week at the hospital. The workshop was planned to have three parts. The duration of the workshop was planned for 1.5 hours, with each part having 25 min and 10 min break between parts. We ended doing only the first two parts in 1.5 hours because more general preparations took time, such as sitting in the right place, explaining the tasks, and showing examples.

The first part consisted of a set of cue cards with possible functionalities for a digital Goal-Plan. Each of the patients' participants had a set of cue cards in different colors. The participants had to read the cue cards individually and chose among the set of cue cards, those cards that involved functionalities that they liked. There were no limitations in the number of cards to choose from. Participants were also given empty cards so they could add more functionalities if they wanted.

As a second task of the first part, participants were asked to choose the five most favorite cue cards. These five cards should be presented later to the other participants in the workshop. After presenting the five cards chosen to the other participants and telling them why the person thought the card was relevant, all participants had to discuss in the group and agree on a set of five cards that all thought were the most important functionalities to have in a digital Goal-Plan that could support them in defining their rehabilitation goals. In the workshop in parallel with the patients' table, we had a table of therapists (staff members) from the DCR doing the same tasks. The set of cue cards that they had was targeted to therapists needed and desired functionalities in a digital Goal-Plan that could support their work in facilitating patient’s rehabilitation.

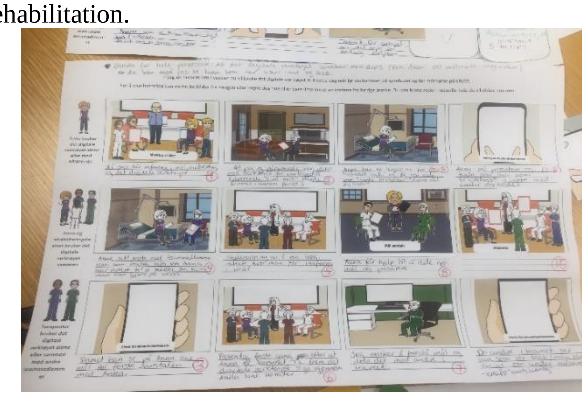


Figure 4. Storyboard created from one of the patient-therapist teams in workshop 5

The second part of the workshop required each of the patient participants to work in pairs with one from the therapists. The task the duo patient-therapist had was to make a story. The story would be on how the digital tool, which had the functionalities the patients and therapists had concluded in the first part as the most desirable ones, would be implemented in the processes at the DCR.

A paper storyboard was given to each of the pairs. The storyboard was divided into three parts to help the patient and the therapist to think about the activities the patient does alone, or activities the therapist does alone or activities they do together. A set of animated pictures picturing an animated fictional patient in specific moments at the hospital, home, or at the outpatient clinic was provided to the patient and the therapist as facilitating cues to make possible the creation of the storyboard. Facilitators told that if participants wanted, they could as well draw, or in case that they wanted to represent some functionalities of the digital tool, they could just paste in the storyboard one of the cue cards of the first session.

The last task was to provide the patient and staff with some choice of technology layouts such as phone, tablet, PC, and some cut out of possible icons and ask them to try to design the interface of the digital solution. However, the time spent in the first two parts did not create the opportunity for doing the last part, so we dropped the idea.

In the workshop, two facilitators participated - the first author and the "knowledgeable third party". Differently from the first case, the "knowledge third party" in this case was not only facilitating the communication but was as well in charge of leading specific sessions together with the first author - designer. The patient workshop was facilitated by the knowledge third party and the staff workshop by the first author. In the second part, each of the facilitators had to lead one of the pairs workshops, sitting on the two opposite corners of the table. We had one therapist more participating in the workshop. Hence, one of the therapists in the second part worked on the tasks alone in order to not shake the balance in the pair's groups.

Workshop 5 had the same outline. This time knowing the limitation in time, we scheduled only the first two parts of the workshop and gave time to other practicalities. The workshop was called "My Interactive Goal-Plan – Owning my rehabilitation". The workshop aimed to discuss how a digital tool can contribute to support the patient to be more in control and involved in her/his rehabilitation process after the patient has defined the goals. 5 patients and 5 therapists from the Sunnaas DCR participated. Two facilitators participated. A third person was involved in supporting materials distribution and making sure that everything was in place while the two facilitators were leading the sessions.

C. *Reflections-on-action*

In this subsection, we present a set of reflections-on-action from workshop outline 3. Reflections are grouped based on the categories borrowed from Hendriks, Truyen, and Duval [9] and also used in Table 1 above. The reflections are presented as insights on best practices or problems we faced based on a look back on our experience while conducting PD with MACI patients. Here we include only reflections that were additional in workshop outline 3. The reflections from workshop outline 1 and 2 have already been taken into consideration before planning workshop outline 3. The following reflections are focused only on the MACI patients. Even though the DCR staff members participated in the workshops together with the patients, their involvement in the process is out of the scope of this paper.

Preparation

- WO3_1 – Dividing the tasks of reDCRitment and planning as described above helped in doing better planning because more time was dedicated to discussing the workshop outline and refining how the tasks in each part would be represented. Moreover, the person in charge of reDCRiting had more time available to dedicate to explain to the patients that had expressed their will to join the workshop, the aim of the workshop, and why their participation would be relevant.
- WO3_2 – The involvement of a knowledgeable third party in the planning of the workshops was relevant for formulating better the workshop aim considering her expertise in the patients' group and the process of rehabilitation. Moreover, her engagement helped in formulating better invitation and workshop description that was suitable for the MACI patients and their challenges in communication.

Tools

- WO3_3 – Adding a written material given to participants during the workshop facilitated the information processing for them. Participants that had difficulties in understanding the requirements in each part of the workshop read what the task was about in the material written as a manual step by step with clear bullet points. The printed material helped them stay focused and have a higher level of understanding of the tasks.
- WO3_4 – Having short sentences and in a simple language suitable for the participants improved understanding and engagement. Moreover, the material was given in Norwegian. This helped the patient understanding and reduced the level of fatigue that speaking and reading on a foreign language can require.
- WO3_5 – Using low fidelity tools like the paper printouts in cue cards or in the animated images made it easier for the MACI patients to contribute to the workshop. The paper tools provided participants with the possibility to rewrite and move around based on their needs. Moreover, there was no fear that a card or a visual image was destroyed as we had a bunch of extra print outs ready to be distributed on needs.

Techniques

- WO3_6 – Cue cards facilitated the process of envisioning a future solution. The patient could agree or not with the hints mentioned in the cue cards. Sometimes an explanation of the cue cards information was needed. In that case, the facilitator would tell a little bit more on what was the aim behind those cards. Having initial cues helped to bring on participants' attention things that they might have forgotten on their own. Moreover, working on the cards and refining the ideas of the cards was expressed from the participants to be easier than having to initiate the thoughts themselves.
- WO3_7 – The openness of the cue cards gave the possibility to the patients to add their personal experiences. Some of the participants gave the cards other meanings based on their understanding and will. Thus, having cue cards not too detailed opened the opportunity for the participants to not just agree with the cues but be able to customize them, as shown in Figure 5.
- WO3_8 – Participants expressed that they found the project relevant and interesting for them and for other patients with MACIs in need or rehabilitation. This was

the main reason they had committed to participate. Moreover, as stated above, most MACI patients are still working, and they are familiar with the notion of workshops. One of the participants was working on service design and was very familiar with the techniques used.

- WO3_9 – During workshop four, considering the time spent on the previous tasks, both facilitators agreed to drop the third part of the workshop. After the workshop, both facilitators expressed that the participants needed a long time to read the cue cards and discussing them. This had created a delay in the previous tasks. Moreover, both facilitators noticed during the workshop, that the tempo of information processing was slower for some participants. Thus, providing them with the time they need is relevant to take into consideration.
- WO3_10 – In the second part of the workshops, we had created a fictional character for the storyboard. In both workshops, we saw that patients' participants were not influenced by this fictional character that we called "Anna". They quite often referred to this character as "me" – "I am the one in the story".
- WO3_11 – The second part of the workshop outline was more demanding than the first one. Participants had to discuss on the cue cards, make sense of them together (patient and therapist) and then make a story. These tasks put a high burden on cognition. However, the usage of the visual cues in the animated form facilitated the envisioning of the future solutions made more concrete in the case of the storyboard. The visual images enhanced creativity and sparked ideas for the story. Patient participants and therapists enjoyed having the visual cues and, as in the case of cue cards, took the freedom to interpret these visual images as they wished. Moreover, in this part, we introduced participants with an example of the storyboard created by the facilitators. Participants had the exemplar as inspiration and did not look at it in detail. However, the exemplar helped them envision what they had to do in the task.
- WO3_12 – Structuring the storyboard and how to build the story helped in making an abstract idea more concrete and the story more approachable for the participants, both patients, and therapists. The duo patient-therapist could divide the activities as instructed in the storyboard. Moreover, the structured way of creating the storyboard served for initiating a discussion on what activities the patients would like or should do alone and as well in which activities during their rehabilitation they can or should interact with the therapist. The structured way of thinking, and building the storyboard facilitated the patients' contribution to the workshop.

Facilitators

- WO3_13 – Coordinating 5 patient participants and 5 therapists as participants required more than two facilitators. In workshop 4, the facilitators had the possibility to sit with the therapist and the patient individually and try to ask in-situ questions when needed. Instead, in workshop 5, both facilitators were moving among groups but were not constantly present while the groups were working. Momentos, in which more investigative questions could have been asked, were lost. These were only realized afterward when hearing the recording.

Participants

- WO3_14 – One patient participant in workshop 4 was tired at the end of the workshop. Instead, all the 5 patient participants in workshop 5 stayed overtime and seemed to enjoy the tasks. From this, we want to highlight that MACI patients capacity variates, and in order to involve everyone and not risk tiring the participants, either participant with the same capacity should be grouped together in workshops, or we should design the workshop based on the capacity of the most fragile participant. This can be established before the workshop while knowing the participants' clinical condition.
- WO3_15 – The patients and the therapists participating in the workshops had a good collaboration. They both collaborated into making the story. Some patients initially struggled in the understating, but the respective therapists supported them by explaining the task so the patients could contribute significantly. The involvement of the staff members as participants in the workshop was not to ask them what the patients need in a digital Goal-Plan (the patients can speak for themselves) but to ask them about their share in the digital solution. Sitting a patient and a therapist together in designing a shared digital solution that will be used by them is not a common practice. Thus, in a future publication, we will expand more on how the collaboration in a PD workshop worked between these two user groups.

Analysis

- WO3_16 – Both facilitators conducted a fast round of reflections-on-action after each of the workshops. Facilitators discussed their individual and common impressions about the workshop and highlighted strong points and downsides in each of the sessions that they were in charge. This was audio recorded for future analysis and reflections. Those immediate reflections-on-action were very helpful in refining the list of reflections presented in this paper because the immediate reflections captured feelings and perceptions, which usually are lost when data is analyzed later in time.

VII. DISCUSSION

In this subsection, we discuss the findings from the reflective analysis from the literature review perspective. We conclude with a list of guidelines for working with people with MACIs for each of the categories initially introduced by Hendriks, Truyen, and Duval [9].

A. Preparation

In our experience, the preparation phase was conducted in close collaboration with domain experts. Thus, the experience that we describe is seen from the perspective of involving domain experts and people experienced with specific patient groups for planning the PD process. The literature recommends getting the consent of the participants at various moments throughout the research process [8]. Our participants did not participate for an extended period in the research. In our workshops, we experienced that the consent prior to the workshop was sufficient. Nevertheless, throughout the workshop, both facilitators were closely observing the participants for signs of fatigue or irritation and informally getting approval that the process was going well for each of the participants. An important insight from our study was the need to provide the information described in the invitation and consent form in different forms (verbal, visual, etc.) to the participants and repeat the informing process many times to ensure that the information is processed, and the person is aware of what s/he is committing to (WO1_3, WO2_1).

Another guideline from the literature is to communicate about the project goals without intermediaries [9] (DG_PP2). In our two cases, the domain experts communicated the project goal to prospective participants. Further, during the workshop, the facilitators repeated the project goal as a precaution to assure that all participants were aware about what they were contributing to. When the intermediaries are people that have knowledge about the cognition challenges of the patient group and are experienced and trained in communicating with them, the intermediaries can be an asset in establishing the communication with the prospect participants and explaining the project goals (WO1_3, WO2_1).

Moreover, in analogy to the literature (DG_PP3), we experienced that there was a need to establish an extra time for general practicalities [9]. However, this time could be managed better if the preparation phase was handled by a group of people who are part of the PD project. Dividing the recruiting and planning process among different persons created more space for ideating better the workshops in workshop outline 3 and as well manage better the recruiting process (WO3_1). The literature states that it is relevant to know the target group well [9][41][42] (DG_PP4), know the patients' deficits so you can adapt to their situation. For researchers and designers, a higher understanding of the MACI patients' cognition challenges can come due to the close collaboration with the domain experts – the rehabilitation specialists working with MACI patients in cognitive rehabilitation. They have deep knowledge about the patient group and can contribute to informing designers. However, PD requires mutual learning and applying this perspective to teach domain experts how designers work can help them provide more knowledge about the patient group (WO1_1). In our case, the selection of the participants was made through the clinic. Assessing abilities through standardized tests [9][4][43] (DG_PP5) was helpful in defining the patients' abilities and disabilities, and for us to plan adequately. They were also useful in the analysis. Ability assessment was not done by the designer but by the healthcare practitioners.

The literature states that it is beneficial to plan and reDCRit participants well in advance (DG_PP6). In the case of MACI patients, we experience that planning well in advance is recommended, especially when the designer leading the PD project is new to working with MACI patients and need to learn more about the patients' needs and situation from the healthcare practitioners. However, the reDCRiting process was done over a short period. This because patients do not stay at the hospital long, and some of them can forget about participation in the workshop if they were reDCRited well in advance. Moreover, patients' condition varies from one day to another (WO1_5). Thus, planning for absent patients is required.

In the preparation phase, we finalized these guidelines for conducting PD with MACI people:

1. Invite the patients and present the information regarding the project in different ways, either text, verbal explanations, images, audio, etc. and make sure to repeat the information several times during the workshops/activities based on the participants' needs.
2. Benefit from the knowledge of domain experts (in this case, the rehabilitation specialists) to recruit and convey the information about the project. They know how to work with MACI people.
3. Plan the PD workshops in collaboration with a multidisciplinary group. Establish mutual learning and make better preparation for the PD process by benefiting from the expertise of everyone.
4. Plan the project well in advance and recruit in a short time. Prepare for absences.

B. Tools

The literature emphasizes the need to involve users in design in appropriate and familiar environments, which take into consideration the deficits of the participants [9][45] [46][47] (DG_T1). The same is true for MACI patients. The hospital environment was familiar, and the participants had previously been in the areas where the workshops took place (WO1_6). Moreover, these areas at the hospital are designed to offer easy accessibility for everyone. Another important element mentioned in the literature is to adapt the language to the participants (DG_T2). In the case of MACI patients, this is extremely relevant. Our reflections from the workshops (WO3_4) and existing literature [19][20] show the importance of using short sentences and an understandable language when addressing MACI patients.

Regarding tools used during workshops as supporting materials for techniques, we found that sometimes using text might be a problem, and it can be more useful to make use of non-verbal elements such as visual stimuli like photos of objects or use physical artifacts [9][50][51][52][53][54]. In our empirical data, we found that having the Goal-Plan in a printed version served as a stimulus for the participants (W02_3). Moreover, we experienced that MACI patients felt more motivated to use workshop tools if these tools were individualized. The MACI patients worked well in manipulating the tools provided initially individually and then sharing the outcome with others.

We did not use contextual cues such as nametags as it has been proposed in the literature [9] (DG_T4). This because the participants coming to the workshops knew each other from before.

In our experience, using low fidelity tools part of the workshop toolkit made it easier for the participants to contribute. However, we lack experience with digital toolkits, and further investigation of conducting PD building on digital toolkits is needed. Despite our lack of experience with digital toolkits, we argue that being aware and considering the fidelity of the toolkit [4] used in a PD project should be a priority. This should be carefully considered with regard to patients' abilities tested through standardized tests.

Finally, in WO1_7 and WO2_2, we highlight how the structuring of the tools became relevant for motivating participants' contributions to the workshops. This is compatible with the rehabilitation theories for building structure in remembering things and focus attention [25] and should be taken into consideration when presenting PD tools in workshops.

Regarding tools, we have the following guidelines for conducting PD with MACI people:

1. Involve users in a familiar environment
2. Use distinctive contextual cues in the toolkit materials
3. Consider the fidelity of the tools in relation to patient-specific cognitive challenges
4. Use a simple language with a positive tone
5. Use visual stimuli which are individually targeted
6. Have clear tools for each part of the workshop and have a structured way of delivering the tools.

C. *Techniques*

Having clear guidelines and techniques for conducting PD with MACI patients that involves a significantly heterogeneous group is difficult. Moreover, techniques can variate based on the technology to be designed. This may put other requirements in place. Here we highlighted insights from our experience within the two projects and five workshops, and we invite other researchers working with MACI patients to refine and supplement the list.

People with cognitive impairments find it challenging to envisioning future solutions [2][9][42][44][48][54]. In the literature, different ways of supporting the envisioning of future solutions are proposed (listed in DG_M1).

In our work with MACI patients, we have found that a task-oriented approach of activities (WO2_4) and narrow scoping of a session (WO2_6) can help the patient to process a line of information at once and to be able to envision more future usage of the solutions. The fear of using the white paper showed the challenge that MACI people have in envisioning a future solution and how the fantasy ability can be undermined when too many options are presented. Thus, as stated in the literature, trying to avoid appealing to the person fantasy and avoid too much choice [8] is adaptable for the MACI patients as well.

A relevant finding influencing the future envisioning is what we called the “teaser of future envisioning” (WO2_8) in the workshop outline 2. The aim is not to ask the participants directly to enter into a fantasy phase but use intermediary tasks that can aid the fantasy of the participants. In the literature is emphasized the relevance of making participants share personal experiences as a start for boosting future envisioning [9][41][44][46][55][56]. The teaser of future envisioning should build on personal experiences that make the participants think about the future.

Another important element for surpassing the challenge in envisioning future solution was the usage of cues in the form of written text cue cards (WO3_6) or cue cards with pictures (WO3_11). The usage of visual cues is recognized in the literature [41][42][55][59].

What we found interesting in our workshops was trying out the power of exemplars as a way to enhance creativity (WO2_9). The usage of examples of designs as a means to aid the fantasy of people with MACIs needs more consideration and further study. However, we can state that it was helpful for our participants who had different aspects of MACIs. It aided their creativity by making them think outside of the box. We observed that the exemplars presented in the form of amateur and not finished designs helped the participants relate more to them and feel more confident in designing themselves as they noticed that no finished and polished designs were expected by them.

In [9], using fictional characters has been defined as useful in envisioning future solutions. However, our participants seemed not to be keen on that. They wanted to be represented and talk about themselves instead of a fictional person. This is also related to rehabilitation theories where patients are motivated to accept and embrace their new selves.

In the literature, providing more hands-on activities and collective prototyping [55][58][59] is seen as contributing to participants' ability to envision a future solution. We experienced that for MACI patients, the envisioning process required a break down into smaller activities that could help the patient create a bigger picture by putting the pieces in each smaller activity together. This is similar to the memory rehabilitation theories [25], which suggest breaking down an activity in smaller steps and train each of the steps slowly, adding one step at the time. Using activities that are familiar is as well helpful to consider in techniques with MACI patients similarly to the findings from the literature [9] [55][59].

Another challenge that people with cognitive impairments face is abstract concepts [41][50][59] (DG_M2). From our cases, we found that MACI patients also have a fear of sketching and the white paper syndrome, hesitating to draw. Based on this, designing more narrowed down (WO2_6) and structured activities (WO3_12) and tell personal stories or personal opinions (WO1_14) can help in surpassing the challenges of MACI patients with abstract concepts.

We also found that people with cognitive impairments are keener on getting involved in designing solutions that are interesting, valuable, and have a real purpose [9][44][63]. We have highlighted the same point in WO3_8 reflection.

Another element to consider in deciding about PD techniques to apply with MACI patients is to provide alternative activities that can support all the participants to engage [9][48][50][65][66] (DG_M7). With the MACI patients, we found that it is important to make an appeal to the individual participants' abilities (WO1_15). Moreover, alternative ways to present the tasks are needed, so it fits the patients' needs. MACI patients experience an increase in the time needed to perform activities. This is called the tempo of performing activities. Adapting to MACI patients' needs in the tempo of activities is very relevant for assuring that patients do not feel overwhelmed and rushed.

In DM_8 we found that it is relevant to consider activities that are flexible and empathic enough to adapt to the needs of the group, for example, activities that can help create a friendly environment [44][46][67], activities that can boost participants self-esteem and confidence [52][68], and activities that can include an element of playfulness [42][52][55]. We experienced that being flexible was required when working with MACI patients. Moreover, serving coffee and biscuits during the breaks helped to create a friendly environment. One of the patients made a video in workshop 5 and shared that with us to express his enthusiasm.

Regarding techniques, we have the following guidelines for conducting PD with MACI people:

1. Having a task-oriented approach where more complicated activities are presented in small steps that build on each other.
2. Having a narrowed scope for the PD sessions and not distracting people with MACIs with general questions.
3. Using cues that can support future envisioning. It is important to consider different ways of presenting the cues. Both text-based and images are useful. The cues should be open so they can offer the possibility for personal interpretations from the participants in the PD workshops.
4. Introduce in workshops “the teaser of future envisioning” and activity that builds on people with MACIs current experiences and ask them to think how these specific experiences can be improved in the future.
5. Take into consideration using exemplars that present examples of what the MACI people are expected to do.
6. Use positive rhetoric when asking for critical opinion. The aim is to not influence MACI people to enter in a negative mindset.
7. Prepare alternative activities that can include all the participants in the workshop independent of their disability.
8. Create a friendly environment by showing empathy and respect toward participants’ experience.
9. Involve MACI people in PD projects that are relevant and interesting for them.
10. Structure the activities as much as possible so it can be easier for the MACI people to conceptualize.
11. Try to avoid fictional characters in the design process. MACI people prefer to refer to themselves in the design.
12. Adapt to the MACI people's tempo while conducting activities in a PD session.
13. Be flexible to changes activities, drop activities, repeat the explanation of activities based on the needs, and the requirements of the MACI people involved in the PD session.

D. *Facilitators*

The literature emphasizes that one of the facilitators' responsibility is to explain clearly the purpose of the events and the role of the participants [9] (DG_F1). Similarly, this was important with the MACI patients were repeating the aim of the event in a clear language, and having it printed out during workshops 4 and 5 helped the participants stay focused and contribute significantly to the workshops. Moreover, the facilitator should try to appeal to the patients' challenges (WO_10) by highlighting that not everyone is perfect [9][69]. The MACI patients all come from a life without their current disability. Thus, making them feel good by emphasizing that the challenges are common among other people without the ABI can break the ice.

The literature also emphasizes that the facilitators should incorporate structure and review in activities [9] [2][43][46][54][59] (DG_F3). They should give time to participants to know each other, have the possibility to repeat, and review parts of the workshops. With MACI patients, this was very relevant. The facilitator should also consider having a slower tempo to adapt to the patients' ability to process information.

Moreover, trying to minimize distraction and keep participants focused [9] (DG_F4) is also a challenging task when working with MACI patients. This can be supported by having more structured and narrowed down workshops where patients have short and clear tasks to perform.

One important finding from our work can be found in WO1_18, the involvement of the “knowledgeable third party” as a facilitator in the workshop. Considering the variations in MACIs, it would be impossible for a designer to be able to have the ability to communicate properly with every variation of cognitive impairment. A person that is specialized for working with MACI patients can support communication. Moreover, in the digitalization project, we saw the knowledgeable third party not only facilitating communication but also leading the PD sessions. This was the result of a long mutual learning process in which the designer and the knowledgeable third party had been involved throughout the “redesign” and “digitalization” projects described above. However, involving the knowledgeable third party and the number of facilitators in a session, in general, should be balanced to the number of participants in order to avoid putting MACI patients in the spotlight.

Regarding Facilitators, we have the following guidelines for conducting PD with MACI patients:

1. Involve a knowledgeable third party as a facilitator for facilitating communication and ultimately leading the sessions.
2. The facilitators should explain clearly the purpose of the events and the role of the participants for each part during the workshop.
3. The facilitators should incorporate a structure in the activities and the review of the activities.
4. The facilitators should enclose some personal information about themselves.
5. The number of facilitators should be balanced with the number of participants so the facilitators can devote more time to each of the participants or participants

groups and ask in-situ questions to uncover meaning in the ideas or provoke new ideas.

E. Participants

Using participant groups with few members is suggested in conducting PD with people with cognitive impairments [2][9][46][52]. This is also true for MACI patients. A number of 4-5 patients were suggested by the domain experts to be a good group size. The duration of the workshops should be short and adapted to the number of participants involved. Thus, enabling everyone to have a say and to not rush the slow tempo of some of the MACI patients. In DG_P1, the guideline is to consider one to one group work in PD sessions with people with cognitive impairments [9][54][71]. Moreover, Yaghoubzadeh, Kramer, Pitsch, and Kopp [70] state that cognitive impairments could be challenging for working in groups. From our reflections (W02_13), we found that participants had experience and worked well in a group. They were able to build on the ideas of others while still keeping their stand if they had a different opinion. The benefit of group work is also compatible with the rehabilitation theories, where group therapies are considered very effective [24][25][84]. However, we stated that not all MACI patients have the same abilities. When working in groups, it should be the facilitator's responsibility to give the same time, attention, and possibility to everyone.

Another guideline from the literature is to pair persons with different deficits into one subgroup [9]. This aims to surpass challenges in individual deficits by working as a group and contributing each with their abilities. In our experience, we noticed that participants who had different cognitive impairments, but the same functioning level could work better in the same group (W0_12).

The literature also recognizes the involvement of caregivers as support in conversation with participants [2][9][41][51][59]. In our case, we had direct caregivers as participants in the same workshop with the patients. Caregivers have usually been involved in the design process as patients' proxies for patients with some forms of cognitive impairments. MACI patients have the capacity to be involved and speak for themselves. Using caregivers as proxies is useful when the user group being represented in not able to be involved. In the MACI patients' this is not the case. In PD with MACI patients, domain experts can support the process of planning the work with patients and make sense of the patients' needs. Meanwhile, MACI patients can participate in PD activities.

In the digitalization project, an MACI patient and a therapist had to work together in making the storyboard. All seven pairs in both workshops 4 and 5 had a good collaboration. Thus, involving in a PD project as participants, both the MACI patients and the therapists in designing digital solutions can result in a positive experience.

Moreover, the literature discusses the elimination of usability problems with the carers of the patients [4][9] and using persons who do not suffer from a deficit to get rid of general design problems [2][9]. In our reflections, we found that the participants in workshop 2 with a milder ABI had the possibility to contribute more in design details.

The literature also highlights the need to involve the kin and the family in the design [73][74] (DG_P6). We have not experienced this in our cases. However, we want to argue that the involvement of the family members or kin should be done only when it is necessary, and the solution designed involves them as well. One of the patients in the workshop, when asked about family involvement, said: "I should decide if I should involve my family".

Regarding Participants, we have the following guidelines for conducting PD with MACI people:

1. Involve participants in group activities where they can work on their own and together with others.
2. Consider a small number of participants for a short period of time.
3. Involve a "knowledgeable third party" to support the conversation with the MACI participants.
4. Use persons with milder cognitive impairments for exploring design details.
5. Promote the involvement of family members as participants only in the design of the solution that involves them and when the MACI person agrees.

F. Analysis

Hendriks, Truyen, and Duval [9] suggest that the researcher should try not to over-analyze the utterance of the participants. Moreover, they suggest being critical to the representativeness of participants. These guidelines are also useful when designing with MACI patients. However, from our experience, we suggest that the reflexive analysis can benefit from the involvement of a team from different disciplines. This can also eliminate the problem of over-analyzing the utterance of the participants because a caregiver can take things less seriously than a designer that is new to the patient group.

Furthermore, implementing a structured reflection on the action right after the workshop where the facilitators reflect on the workshop in general, tools, techniques, participants, and their behavior can be very relevant to the analysis later because it captures the feelings at the moment, which often can pass undocumented.

Regarding Analysis, we have the following guidelines for conducting PD with MACI people:

1. Try not to over-analyze the utterance of the participants.
2. Have a critical attitude toward the representativeness of participants.
3. Involve people from different disciplines in the analysis, especially rehabilitation specialist in MACIs.
4. Incorporate a reflection-on-action structure among the facilitators right after the workshop.

VIII. CONCLUSION

In this paper, we present a set of guidelines for researchers and designers to conduct PD with people with MACIs. We have initially presented in Table 1 a summary of guidelines drawn from previous studies of conducting PD with people with cognitive impairments. Then, we have presented two PD projects that we conducted with MACI patients and presented a set of reflections from each of the workshop outlines we have been working with. The reflections have been further discussed in regard to the existing literature, and finally, a set of guidelines for conducting PD with MACI people has been introduced. While the guidelines are the final outcome of the paper, the rich description of the reflections-on-action is also a contribution to PD, which put emphasis on the situated knowledge generated in PD workshops. These rich descriptions in some cases are even more relevant to PD because they represent a story derived from the experience of the PD researcher that has conducted the study and highlight things that are usually overlooked on more formal guidelines. Hendriks, Slegers and Duysburgh [85] state that a good way to go forward on a codesign approach for people suffering from some form of impairments is "*facilitating researchers and designers to share experiences, best practices, lessons learned, and so on …in the form of method stories*"

People with MACIs compound a significant part of our society. This is increasing with the increase in the tempo of life. People are more in danger of accidents and consequently are at risk of having more accidental brain damage. In people with MACI in many cases, there are no physical impairments. MACI people work, go to school and try to live their life to the fullest. However, their daily life is challenging due to fatigue, memory problems, attention problems, loss of executive functioning, etc. Thus, they need to adapt their lifestyle to their new self and make use of aids to keep up with daily life activities. Technology can help in assisting MACI people.

In designing these new technologies for them, we need to involve the MACI people in design. They can significantly contribute to the design if the right means for enabling their contribution are provided. That is what we want to achieve with this paper.

We contribute by giving PD practitioners a list of guidelines for working with MACI people. Moreover, through these guidelines, we aim to make technologists turn attention to the MACI people and design more supportive technologies for them.

The number of participants involved in our study is small in comparison to the heterogeneity of the MACI people group. Our guidelines are not a final list, and we hope that more researchers will investigate on this user group and expand our lists. These guidelines are in the form of recommendation, and they should be combined based on the situation at hand, in which PD researchers and designers critically reflect on what can be adapted in their specific case and what not and what is the consequence in the PD process if one of the guidelines is not taken in consideration.

In the future, we will continue testing our guidelines in further projects with this user group. Additionally, we want to investigate how to involve more digital tools in designing together with MACI people and how we can involve in the best way possible the MACI people in the co-development of different types of digital tools meant for them, besides the cases presented in this paper. Furthermore, we want to investigate how much are the MACI people willing to participate in PD practices, and where do we, as researchers and designers draw the line.

ACKNOWLEDGMENTS

We are grateful to all those patients at Sunnaas rehabilitation hospital that participated in our workshops. Their participation made this publication possible. Moreover, we want to deeply thank all our partners at the hospital, specifically the Department of Cognitive Rehabilitation that worked with us in building the PD process and who share the same values as we do in terms of involving the patients in designing technologies or services meant for them. The activities of the case 2 presented in this paper were financed by the INTERREG Sverige-Norge project "VITAL - för den goda hälsan", under the grant number: 20202391.

REFERENCES

[1] K. Çarçani and H. Holone, "Participatory Design "Method Story": The Case of Patients Living with Mild Acquired Cognitive Impairments," *ACHI 2019, The Twelfth International Conference on Advances in Computer-Human Interactions*, pp. 210-217, 2019.

[2] S. Lindsay, K. Brittain, D. Jackson, C. Ladha, K. Ladha, and P. Olivier, "Empathy, participatory design and people with dementia," *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 521-530.

[3] J. McGrenere, R. Davies, L. Findlater, P. Graf, M. Klawe, K. Moffatt, B. Purves, and S. Yang, "Insights from the aphasia project: designing technology for and with people who have aphasia," *SIGCAPH Comput. Phys. Handicap.*, no. 73-74, pp. 112-118, 2002.

[4] K. Moffatt, J. McGrenere, B. Purves, and M. Klawe, "The participatory design of a sound and image enhanced daily planner for people with aphasia," *In Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 407-414.

[5] A. Eghdam, "Understanding how persons with mild acquired cognitive impairment use and experience information and communication technology: an exploratory study," 2016.

[6] C. A. Taylor, J. M. Bell, M. J. Breiding, and L. Xu, "Traumatic brain injury–related emergency department visits, hospitalizations, and deaths—United States, 2007 and 2013," *MMWR Surveillance Summaries,* vol. 66, no. 9, pp. 1, 2017.

[7] A. Eghdam, J. Scholl, A. Bartfai, and S. Koch, "Information and communication technology to support self-management of patients with mild acquired cognitive impairments: systematic review," *Journal of medical internet research,* vol. 14, no. 6, pp. e159, 2012.

[8] J. Simonsen and T. Robertson, "Participatory Design: an introduction," *Routledge international handbook of participatory design*, pp. 21-38: Routledge, 2012.

[9] N. Hendriks, F. Truyen, and E. Duval, "Designing with dementia: Guidelines for participatory design together with persons with dementia," *In: IFIP Conference on Human-Computer Interaction,* pp. 649-666.

[10] D. A. Schön, "The reflective practitioner: How professionals think in action," *Routledge*, 2017.

[11] J. Toglia, “Cognitive-perceptual retraining and rehabilitation,” *Willard and Spackman's occupational therapy*, pp. 607-629, 2003.

[12] K. D. Cicerone, L. C. Smith, W. Ellmo, H. R. Mangel, P. Nelson, R. F. Chase, and K. Kalmar, “Neuropsychological rehabilitation of mild traumatic brain injury,” *Brain Injury,* vol. 10, no. 4, pp. 277-286, 1996.

[13] L. Turner-Stokes, “Evidence for the effectiveness of multi-disciplinary rehabilitation following acquired brain injury: a synthesis of two systematic approaches,” *Journal of rehabilitation Medicine,* vol. 40, no. 9, pp. 691-701, 2008.

[14] C. Nilsson, A. Bartfai, and M. Löfgren, “Holistic group rehabilitation–a short cut to adaptation to the new life after mild acquired brain injury,” *Disability and rehabilitation,* vol. 33, no. 12, pp. 969-978, 2011.

[15] J. Toglia, “A dynamic interactional approach to cognitive rehabilitation,” *Cognition and occupation across the life span: Models for intervention in occupational therapy,* vol. 2, pp. 29-72, 2005.

[16] G. P. Prigatano, “Principles of neuropsychological rehabilitation,*” Oxford University Press*, 1999.

[17] J. M. Krogstad, “Hva er ervervet hjerneskade?,” *Konsis Grafisk AS*, 2011.

[18] G. E. Carlsson, A. Möller, and C. Blomstrand, “Managing an everyday life of uncertainty–a qualitative study of coping in persons with mild stroke,” *Disability and rehabilitation,* vol. 31, no. 10, pp. 773-782, 2009.

[19] Sunnaas, “Hverdagsliv med kognitive endringer: Praktiske råd for deg som lever hverdagslivet med kognitive endringer,” *Sunnaas Sykehus HF*, Flisa Trykkeri AS, 2011.

[20] Sunnaas, “Sosialt liv med kognitive endringer: Praktiske råd for deg som lever livet med kognitive endringer,” *Sunnaas Sykehus HF*, Flisa Trykkeri AS, 2011.

[21] R. M. Ruff, G. L. Iverson, J. T. Barth, S. S. Bush, D. K. Broshek, N. Policy, and P. Committee, “Recommendations for diagnosing a mild traumatic brain injury: a National Academy of Neuropsychology education paper,” *Archives of Clinical Neuropsychology,* vol. 24, no. 1, pp. 3-10, 2009.

[22] F. Becker, M. Kirmess, S. Tornås, and M. Løvstad, “A description of cognitive rehabilitation at Sunnaas Rehabilitation Hospital–balancing comprehensive holistic rehabilitation and retraining of specific functional domains,” *NeuroRehabilitation,* vol. 34, no. 1, pp. 87-100, 2014.

[23] E. C. Haskins, K. D. Cicerone, and L. E. Trexler, “Cognitive rehabilitation manual: Translating evidence-based recommendations into practic,” *ACRM Publishing*, 2012.

[24] B. A. Wilson, F. Gracey, J. J. Evans, and A. Bateman, “Neuropsychological rehabilitation: Theory, models, therapy and outcome,” *Cambridge University Press*, 2009.

[25] K. D. Cicerone, C. Dahlberg, K. Kalmar, D. M. Langenbahn, J. F. Malec, T. F. Bergquist, T. Felicetti, J. T. Giacino, J. P. Harley, D. E. Harrington, J. Herzog, S. Kneipp, L. Laatsch, and P. A. Morse, “Evidence-based cognitive rehabilitation: Recommendations for clinical practice,” *Archives of Physical Medicine and Rehabilitation,* vol. 81, no. 12, pp. 1596-1615, 2000/12/01/, 2000.

[26] G. Bjerknes and T. Bratteteig, “User participation and democracy: A discussion of Scandinavian research on system development,” *Scandinavian Journal of information systems,* vol. 7, no. 1, pp. 1, 1995.

[27] T. Robertson and I. Wagner, “Engagement, representation and politics-in-action,” *The Handbook of Participatory Design, edited by J. Simonsen and T. Robertson*, pp. 64-85, 2012.

[28] E. Björgvinsson, P. Ehn, and P.-A. Hillgren, "Participatory design and democratizing innovation." pp. 41-50.

[29] J. Greenbaum and F. Kensing, "Heritage: having a say," *Routledge International Handbook of Participatory Design*, pp. 41-56: Routledge, 2012.

[30] E. Brandt, T. Binder, and E. B.-N. Sanders, "Tools and techniques: ways to engage telling, making and enacting," *Routledge international handbook of participatory design*, pp. 165-201: Routledge, 2012.

[31] D. Corring, R. Campbell, and A. Rudnick, “Cognitive remediation for inpatients with schizophrenia or schizoaffective disorder using "smart" technology in a simulated apartment: a feasibility and exploratory study,” *in Proceedings of the 9th international conference on Toward useful services for elderly and people with disabilities: smart homes and health telematics*, Montreal, Canada, 2011, pp. 286-289.

[32] A. L. Culén and M. van der Velden, "The digital life of vulnerable users: designing with children, patients, and elderly," *In Scandinavian Conference on Information Systems*, pp. 53-71.

[33] N. Hendriks, K. Slegers, and P. Duysburgh, “Codesign with people living with cognitive or sensory impairments: a case for method stories and uniqueness,” *CoDesign,* vol. 11, no. 1, pp. 70-82, 2015.

[34] R. B. Rosseland, “Design and evaluation of an interactive music system for exercise and physical activity with Alzheimer's patients,” *SoundEffects-An Interdisciplinary Journal of Sound and Sound Experience,* vol. 6, no. 1, pp. 4-22, 2016.

[35] I. Nakarada-Kordic, N. Hayes, S. D. Reay, C. Corbet, and A. Chan, “Co-designing for mental health: creative methods to engage young people experiencing psychosis,” *Design for Health,* vol. 1, no. 2, pp. 229-244, 2017/07/03, 2017.

[36] M. Augstein, T. Neumayr, R. Ruckser-Scherb, and S. Dielacher, "Collaboration Around an Interactive Tabletop in Rehabilitation Settings," *Collaboration Meets Interactive Spaces*, pp. 425-442: Springer, 2016.

[37] M. Balaam, S. Rennick Egglestone, G. Fitzpatrick, T. Rodden, A.-M. Hughes, A. Wilkinson, T. Nind, L. Axelrod, E. Harris, and I. Ricketts, "Motivating mobility: designing for lived motivation in stroke rehabilitation," *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 3073-3082.

[38] A. L. Threatt, J. Merino, K. E. Green, I. Walker, J. O. Brooks, and S. Healy, "An assistive robotic table for older and post-stroke adults: results from participatory design and evaluation activities with clinical staff," *In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 673-682.

[39] A. L. Faria and I. B. Sergi Bermúdez, "Development and evaluation of a web-based cognitive task generator for personalized cognitive training: a proof of concept study with stroke patients," *in Proceedings of the 3rd 2015 Workshop on ICTs for improving Patients Rehabilitation Research Techniques*, Lisbon, Portugal, 2015, pp. 1-4.

[40] C. Magnusson, H. A. Caltenco, D. McGookin, M. Kytö, I. Hjaltadóttir, T. B. Hafsteinsdóttir, H. Jónsdóttir, and I. Bjartmarz, "Tangible Interaction for Stroke Survivors: Design Recommendations," *in Proceedings of the Eleventh International Conference on Tangible, Embedded, and Embodied Interaction*, Yokohama, Japan, 2017, pp. 597-602.

[41] J. Galliers, S. Wilson, A. Roper, N. Cocks, J. Marshall, S. Muscroft, and T. Pring, "Words are not enough: empowering people with aphasia in the design process." pp. 51-60.

[42] J. M. Mayer and J. Zach, "Lessons learned from participatory design with and for people with dementia," *in Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services*, Munich, Germany, 2013, pp. 540-545.

[43] M. Wu, B. Richards, and R. Baecker, "Participatory design with individuals who have amnesia," *in Proceedings of the eighth conference on Participatory design: Artful integration: interweaving media, materials and practices - Volume 1*, Toronto, Ontario, Canada, 2004, pp. 214-223.

[44] S. Lindsay, D. Jackson, G. Schofield, and P. Olivier, "Engaging older people using participatory design," *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 1199-1208.

[45] E. Grönvall and M. Kyng, "Beyond Utopia: reflections on participatory design in home-based healthcare with weak users," *Proceedings of the 29th Annual European Conference on Cognitive Ergonomics,* pp. 189-196.

[46] S. Frennert, B. Östlund, and H. Eftring, "Capturing seniors' requirements for assistive robots by the use of attention cards," *in Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design*, Copenhagen, Denmark, 2012, pp. 783-784.

[47] J. Wherton and D. Prendergast, "The building bridges project: involving older adults in the design of a communication technology to support peer-to-peer social engagement," *Symposium of the Austrian HCI and Usability Engineering Group,* pp. 111-134.

[48] R. M. Branco, J. Quental, and Ó. Ribeiro, "Playing with personalisation and openness in a codesign project involving people with dementia," *Proceedings of the 14th Participatory Design Conference*, pp. 61-70.

[49] A. Martin-Hammond, S. Vemireddy, and K. Rao, "Engaging Older Adults in the Participatory Design of Intelligent Health Search Tools," *in Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare*, New York, NY, USA, 2018, pp. 280-284.

[50] S. K. Kane, B. Linam-Church, K. Althoff, and D. McCall, "What we talk about: designing a context-aware communication tool for people with aphasia," *in Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility*, Boulder, Colorado, USA, 2012, pp. 49-56.

[51] L. C. L. d. F. Borges, L. Filgueiras, C. Maciel, and V. Pereira, "A customized mobile application for a cerebral palsy user," *in Proceedings of the 31st ACM international conference on Design of communication*, Greenville, North Carolina, USA, 2013, pp. 7-16.

[52] S. Baker, J. Waycott, F. Vetere, and T. Hoang, "The technology explorers: Partnering with older adults to engage with virtual reality and virtual avatars," *Ageing and Digital Technology*, pp. 231-246: Springer, 2019.

[53] O. K. Richards, "Exploring the Empowerment of Older Adult Creative Groups Using Maker Technology," *in Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, Denver, Colorado, USA, 2017, pp. 166-171.

[54] M. Massimi, R. M. Baecker, and M. Wu, "Using participatory activities with seniors to critique, build, and evaluate mobile phones," *Proceedings of the 9th international ACM SIGACCESS conference on Computers and accessibility,* pp. 155-162.

[55] V. Orso, A. Spagnolli, L. Gamberini, F. Ibañez, and M. E. Fabregat, "Involving Older Adults in Designing Interactive Technology: The Case of SeniorChannel," *in Proceedings of the 11th Biannual Conference on Italian SIGCHI Chapter*, Rome, Italy, 2015, pp. 102-109.

[56] A. Talamo, M. Camilli, L. Di Lucchio, and S. Ventura, "Information from the past: how elderly people orchestrate presences, memories and technologies at home," *Universal Access in the Information Society,* vol. 16, no. 3, pp. 739-753, 2017.

[57] W. Kopec, R. Nielek, and A. Wierzbicki, "Guidelines toward Better Participation of Older Adults in Software Development Processes Using a New SPIRAL Method and Participatory Approach," *Proceedings of the 11th International Workshop on Cooperative and Human Aspects of Software Engineering,* pp. 49-56.

[58] R. J. Wirth, C. W. Yuan, B. V. Hanrahan, J. M. Carroll, M. B. Rosson, and J. Bindá, "Exploring Interactive Surface Designs for Eliciting Social Activity from Elderly Adults," in *Proceedings of the 2016 ACM International Conference on Interactive Surfaces and Spaces*, Niagara Falls, Ontario, Canada, 2016, pp. 403-408.

[59] L. M. Muriana and H. Hornung, "Towards Participatory Prototyping with Older Adults with and without Cognitive

Impairment: Challenges and Lessons Learned," *IFIP Conference on Human-Computer Interaction*, pp. 344-363.

[60] A. Dearden, A. Lauener, F. Slack, C. Roast, and S. Cassidy, "Make it so! Jean-Luc Picard, Bart Simpson and the design of e-public services," *Proceedings of the ninth conference on Participatory design: Expanding boundaries in design-Volume 1*, pp. 67-76.

[61] K. Hara, C. Chan, and J. E. Froehlich, "The Design of Assistive Location-based Technologies for People with Ambulatory Disabilities: A Formative Study," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, San Jose, California, USA, 2016, pp. 1757-1768.

[62] R. Davies, S. Marcella, J. McGrenere, and B. Purves, "The ethnographically informed participatory design of a PD application to support communication," *ACM SIGACCESS Accessibility and Computing*, no. 77-78, pp. 153-160, 2003.

[63] W. Kopec, K. Skorupska, A. Jaskulska, K. Abramczuk, R. Nielek, and A. Wierzbicki, "LivingLab PJAIT: towards better urban participation of seniors," in *Proceedings of the International Conference on Web Intelligence*, Leipzig, Germany, 2017, pp. 1085-1092.

[64] K. Rodil, M. Rehm, and A. L. Krummheuer, "Co-designing social robots with cognitively impaired citizens," in *Proceedings of the 10th Nordic Conference on Human-Computer Interaction*, Oslo, Norway, 2018, pp. 686-690.

[65] A. Hornof, H. Whitman, M. Sutherland, S. Gerendasy, and J. McGrenere, "Designing for the "Universe of One": Personalized Interactive Media Systems for People with the Severe Cognitive Impairment Associated with Rett Syndrome," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, Denver, Colorado, USA, 2017, pp. 2137-2148.

[66] K. Slegers, N. Hendriks, P. Duysburgh, R. Maldonado Branco, B. Vandenberghe, and E. Brandt, "Sharing Methods for Involving People with Impairments in Design: Exploring the Method Story Approach," *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems,* pp. 3331-3338.

[67] M. Zancanaro, S. Gabrielli, A. Jameson, C. Leonardi, E. Not, and F. Pianesi, "Virtual helper or virtual card player? Contrasting responses of older users," *Your Virtual Butler*, pp. 70-78: Springer, 2013.

[68] W. Kopeć, B. Balcerzak, R. Nielek, G. Kowalik, A. Wierzbicki, and F. Casati, "Older adults and hackathons: a qualitative study," *Empirical Software Engineering,* vol. 23, no. 4, pp. 1895-1930, August 01, 2018.

[69] M. Wu, R. Baecker, and B. Richards, "Participatory design of an orientation aid for amnesics," *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 511-520.

[70] R. Yaghoubzadeh, M. Kramer, K. Pitsch, and S. Kopp, "Virtual agents as daily assistants for elderly or cognitively impaired people," *International workshop on intelligent virtual agents*, pp. 79-91.

[71] V. A. Vanden Abeele and V. Van Rompaey, "Introducing human-centered research to game design: designing game concepts for and with senior citizens," *CHI'06 extended abstracts on Human factors in computing systems,* pp. 1469-1474.

[72] I. Colonius, S. Budde, and R. Annicchiarico, "Participatory design for challenging user groups: a case study," *Proceedings of the 28th Annual European Conference on Cognitive Ergonomics,* pp. 339-340.

[73] N. Hendriks, L. Huybrechts, A. Wilkinson, and K. Slegers, "Challenges in doing participatory design with people with dementia," *Proceedings of the 13th Participatory Design Conference: Short Papers, Industry Cases, Workshop Descriptions, Doctoral Consortium papers, and Keynote abstracts-Volume 2*, pp. 33-36.

[74] K. Slegers, A. Wilkinson, and N. Hendriks, "Active collaboration in healthcare design: participatory design to develop a dementia care app," *in CHI '13 Extended Abstracts on Human Factors in Computing Systems*, Paris, France, 2013, pp. 475-480.

[75] K. Slegers, P. Duysburgh, and N. Hendriks, "Participatory design with people living with cognitive or sensory impairments," *in CHI '14 Extended Abstracts on Human Factors in Computing Systems*, Toronto, Ontario, Canada, 2014, pp. 49-52.

[76] J. McGrenere, R. Davies, L. Findlater, P. Graf, M. Klawe, K. Moffatt, B. Purves, and S. Yang, "Insights from the aphasia project: designing technology for and with people who have aphasia," *Proceedings of the 2003 conference on Universal usability*, pp. 112-118.

[77] J. Lazar, C. Woglom, J. Chung, A. Schwartz, Y. G. Hsieh, R. Moore, D. Crowley, and B. Skotko, "Co-design process of a smart phone app to help people with down syndrome manage their nutritional habits," *Journal of usability studies*, vol. 13, no. 2, pp. 73-93, 2018.

[78] J. Löwgren and E. Stolterman, "Thoughtful interaction design: A design perspective on information technology," *Mit Press*, 2004.

[79] D. A. Schön, "Educating the reflective practitioner," Jossey-Bass San Francisco, 1987.

[80] R. Jungk and N. Müllert, "Future Workshops: How to create desirable futures," *Institute for Social Inventions London*, 1987.

[81] W. Gaver and J. Bowers, "Annotated portfolios," *Interactions,* vol. 19, no. 4, pp. 40-49, 2012.

[82] W. Gaver, "What should we expect from research through design?," *Proceedings of the SIGCHI conference on human factors in computing systems,* pp. 937-946.

[83] W. W. Gaver, J. Bowers, A. Boucher, H. Gellerson, S. Pennington, A. Schmidt, A. Steed, N. Villars, and B. Walker, "The drift table: designing for ludic engagement," *CHI'04 extended abstracts on Human factors in computing systems*, pp. 885-900.

[84] C. Wilson, L. Sitbon, M. Brereton, D. Johnson, and S. Koplick, "'Put yourself in the picture': designing for futures with young adults with intellectual disability," in *Proceedings of the 28th Australian Conference on Computer-Human Interaction*, Launceston, Tasmania, Australia, 2016, pp. 271-281.

[85] N. Hendriks, K. Slegers, and P. Duysburgh, "Codesign with people living with cognitive or sensory impairments: a case for method stories and uniqueness," *CoDesign*, vol. 11, no. 1, pp. 70-82, 2015.

Hazard Notifications Around a Vehicle Using Seat Actuators

Akimasa Suzuki, Kaoru Horie, Satoru Otobe and Yoshitoshi Murata
Department of Information Science,
Iwate Prefectural University
020-0693, Sugo, Japan
suzuki_a@iwate-pu.ac.jp

Shoma Fujimura
Ad-Sol Nissin Corporation,
Kawasaki-shi, 210-0804,
Kanagawa, Japan
oaurn85ns@outlook.jp

Abstract—**This paper examines the robustness of our proposed haptic notification system against the different types and layers used for driving seat cushions. While many car manufacturers provide useful side and rear collision warning systems with sound alarms or visual monitors, the addition of similar notifications can confuse a driver because they already need to be aware of many visual targets such as mirrors, monitors, and environmental sounds. Therefore, we have investigated a haptic notification system that uses the driver's buttocks. The results show that drivers can correctly identify the directions of five vibrating motors, three intensity settings, and three obstacle types (*i.e.*, pedestrians, vehicles, and motorcycles). In this paper, we investigate whether drivers can discriminate the direction, intensity of vibrations, and vibration patterns of the system through their buttocks to identify the obstacle direction, degree of risk, and the type of obstacle, even if the vibrations are attenuated by the seat cushion. The results indicate the high potential of the haptic sensation system to notify the driver of obstacles, especially those located in the blind spot.**

Keywords-Vibro-Tactile Notification; Type of Obstacle; Buttocks; Acoustic Haptic Actuators; Seat Cushion.

I. INTRODUCTION

This paper is an extension of the work initially presented in the The Ninth International Conference on Advances in Vehicular Systems, Technologies and Applications [1].

There has been considerable research in investigating accident prevention systems for vehicles, particularly in relation to developing driving support systems that will transition to autonomous driving systems. However, to realize autonomous driving systems, we must overcome problems related to cyber-security measures and traffic laws (*e.g.*, responsibility for accidents by autonomous cars [2]), which could take time. Additionally, as many people enjoy driving, the demand for manual driving as a hobby is unlikely to fade. Driving support systems will thus remain an important feature. Moreover, despite the high number of driving support systems used in Japan, many fatal vehicle accidents are caused by violations of safe driving practices, such as failing to keep eyes on the road, careless driving, and failing to make safety checks [3], thus highlighting the need to develop more techniques that support drivers.

To develop a support system that helps drivers to avoid vehicle accidents, the system needs to quickly and accurately sense information and notify the driver so that he/she can make a rapid judgement. Most car manufacturers now install highly accurate sensor systems at the front and rear of their vehicles at a low cost. Support systems located at the front of a vehicle use vision [4] or radar [5] sensors to prevent careless driving and overcome a driver's failure to make safety checks, while support systems located at the rear of a vehicle use sensors and notification systems to monitor a driver's rear view and blind spot [6]. These systems use sound or visual images to alert drivers to potential hazards.

Visual images can quickly notify a driver about many kinds of information using shapes and colors. As vision is the dominant human sense [7], many notifications rely on the driver's vision, including the front view, mirror, tachometer, speedometer, navigation system, and indicators. There is a concern, therefore, that excessive visual information could affect a driver's capacity to adhere to safe driving practices [3]. We thus consider that developing an additional visual notification may cause the driver to confuse it with conventional visual notifications.

Many conventional systems also provide information to drivers in the form of sounds (*e.g.*, alerts by horn; car audio, including radio; and alarms for reverse gear, pre-collision, and lane departure). Directions presented by a satellite navigation system are also expressed through the vehicle's stereo system. To avoid confusing the driver, we considered creating different sounds, pitches, and patterns for each type of obstacle; however, these would not be intuitive. Additionally, notifying a driver using speech would be too slow to get communicate the message in time. It is also difficult to apply a system using sound on a late-night bus travelling long distances because sleeping passengers may get up by the alert.

Therefore, we proposed a system that uses haptic sensations to quickly notify drivers of possible hazards or obstacles surrounding the vehicle [8]. Our proposed system has higher immediacy and directional resolution than notifications using sounds. As no driver notification system currently uses haptic sensations, we do not have to consider conflicts in this area. Our proposed notification system uses vibro-tactile haptic devices that remain in constant contact with the driver's buttocks. We evaluated the system's robustness against cushion type for determining the direction and intensity of vibrations and road conditions. A high intensity expresses the extent of the danger and the direction of the vibration indicates the location of the hazard. The system can also alert the driver to different types of obstacles, such as a pedestrian, car, or bike. The results indicated a high potential for notifying drivers of obstacles, especially those located in the blind spot.

To support safe car driving, our proposed haptic notification system installing vibration alerts into a driving seat. This paper examines its robustness against different types and layers used for driving seat cushions. We also discussed vibrating waveforms by using a real 4-wheel vehicle, and illuminated higher accuracy on notification by using category-deformed and normalized waves between each actuator than real and certain vibration waves, respectively. Moreover, we discussed notifying accuracies of our system from an experiment combined with conventional notification system with a display and a sound speaker. From this experiment, we also discussed whether our system is obstructive or useful to driving.

The remainder of the paper is structured as follows. Section II discusses relevant studies, Section III describes the proposed system, Section IV describes the modulated waveforms generated for precise notification, Section V presents the experiments to test the robustness of the system, Section VI describes the system mounted on a real 4-wheel vehicle, Section VII describes normalization by using deformed waveform, Section VIII describes the discussion for combination with conventional visual and sound alert system, and Section IX presents our conclusions.

II. Relevant Studies

Many practical driving support systems apply image sensors [4], radar [5], and ultrasonic sensors [9] to detect pedestrians and other vehicles with high accuracy. Around-view monitors are increasingly being used for automatic parking [10] and lane-detection systems are being applied using three-dimensional (3-D) laser imaging detection and ranging (LIDAR) [11]. Despite their weakness to other noise sound, ultrasonic sensors can now be installed in driving support systems for a low cost, while the cost of 3-D LIDAR is also dropping. These devices can be used to detect not only the presence of an obstacle, but also the type of obstacle (*e.g.*, pedestrian, vehicle, or motorcycle). However, in this research, we focus on creating a notification method to alert drivers to the potential hazards, rather than the development of a sensor system.

Previous research on evaluating seat comfort has demonstrated that buttocks are sensitive to tactile sensations [12]. Although not used in the driving seat, some studies have reported the effectiveness of vibro-tactile devices for notifying drivers of directions when using a wearable device such as a belt [13]. The directions of obstacles could be detected by using vibro-tactile devices on the seatback [14] because the back is more sensitive than the buttocks; however, as drivers need to lean against the backrest, the system might have a negative effect on the driver's posture. A vehicle notification device using vibro-tactile devices on the buttocks was therefore developed [15], although the system was unable to indicate the direction of a hazard to the driver.

In a gaming device, vibro-tactile devices are used to link a virtual object with reality [16]. Therefore, we consider applying a vibro-tactile device to notify the driver of essential information related to potential hazards based on the intensity and direction of vibrations. Tactile sensations can include rubbing, pain, pressure, and warmth. On the streets, tapping on the shoulder is a popular method for pedestrians to alert each other. To our knowledge, this study is the first to notify a driver of information such as the direction of a hazard in relation to the vehicle, the extent of the urgency based on the intensity of the vibration, and the type of hazard by the vibration pattern expressed using vibro-tactile devices located below the buttocks.

III. Vibro-tactile Notification System Around a Vehicle Using Seat Actuators

We will indicate our proposed vibro-tactile notification system in this section.

A. System Architecture

For our proposed system, we utilized a vibrating motor with an ACOUSTICHAPTIC™ actuator developed by Foster Electric Company Limited. The acoustic haptic actuator is a kind of woofer that comes into direct contact with the driver's buttocks. Fig. 1 shows the hardware layout for this system. The edited waves were played on a PC, and the five actuators vibrated on the seat, as shown in Fig. 1. These actuators contact with the back of the driver's knees. We used the AP05 amplifier produced by Fostex.

In this experiment, we administered four vibrating patterns of the same intensity, representing different obstacle types, to fifteen participants. The participants were asked to identify the type of obstacle from the vibration pattern. We conducted five trials in a random order for each participant. The vibrations included the sound of footsteps from leather shoes [17], the sound of a V6 engine revving up [18], the sound of a bus driving uphill [19], an idling sound [20] as obstacle types of pedestrian, small and large four-wheeled vehicles, and a motorcycle, respectively. We hypothesized that drivers would intuitively recognize the type from the vibration pattern of the real sound.

As shown in Fig. 1, up to three layers of urethane cushions were placed over the actuator to evaluate the robustness. The thickness of each layer is 2 cm. We also define "layer 0" to mean that nothing is placed over the actuators. We utilize three types of urethane cushion, with specifications shown in Table I. We defined 20 ss, 35 s, and BZ-10 constructed by Toyo Quality One as soft, highly resilient, and less resilient cushions, respectively, as shown in Table I.

We generated vibration waveforms from these sound data, to decreasing up to 2 kHz and increasing between 55 Hz and 110 Hz, which are resonance frequencies of the ACOUSTICHAPTIC™ actuator. Fig. 2 shows the waveforms of these four vibrating patterns, *i.e.*, (a) a pedestrian, (b) a motorcycle, (c) a small 4-wheel vehicle, and

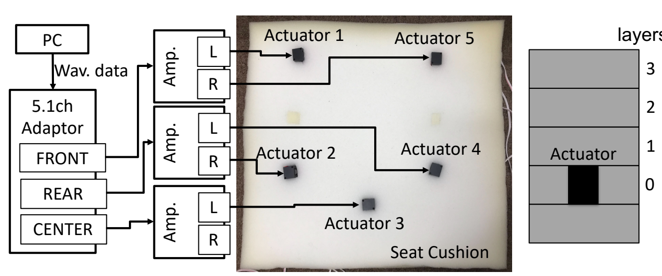


Figure 1. Hardware layout of the notification system using an ACOUSTICHAPTIC™ actuator.

TABLE I. CUSHION MATERIALS FOR EVALUATION.

	Soft	**Highly Resilient**	**Less Resilient**
Constructor	Toyo Quality One	Toyo Quality One	Toyo Quality One
Product name	20 ss	35 s	BZ-10
Density ($\mathrm{kg/m^3}$)	20 ± 2	55 ± 2	35 ± 3
Hardness (N)	30 ± 15	45 ± 15	60 ± 15
Tensile intensity (kPa)	$50 \leq$	$60 \leq$	$30 \leq$
Elongation (%)	$200 \leq$	$100 \leq$	$80 \leq$
Tensile intensity (N/cm)	$3.0 \leq$	$2.0 \leq$	$2.0 \leq$
Compressive residual strain (%)	$10 \geq$	$12 \geq$	$15 \geq$

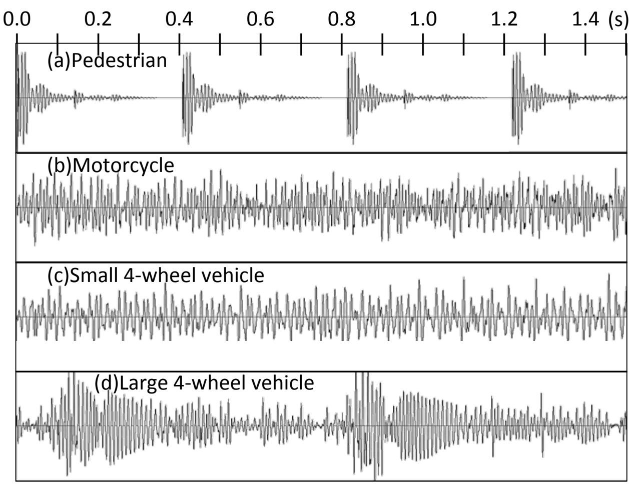


Figure 2. Waveforms of the vibrations for the four obstacle types.

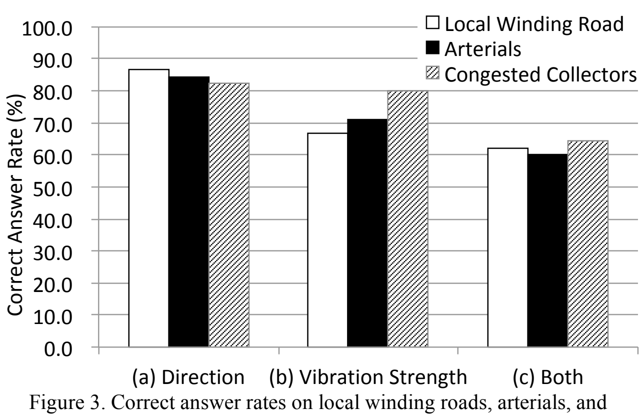


Figure 3. Correct answer rates on local winding roads, arterials, and congested collectors.

(d) a large 4-wheel vehicle. The horizontal and vertical axes indicate the time and amplitude, respectively. A waveform of the pulse vibration with a walking frequency of 0.4-s intervals was used for the pedestrian. We applied the 55-Hz and 110-Hz resonance frequencies to large and small four-wheeled vehicles, respectively. The amplitude of the waveforms was normalized because we utilized the different amplitudes (*i.e.*, the intensity of the vibration) to express the urgency of the degree of risk or the distance to the obstacles.

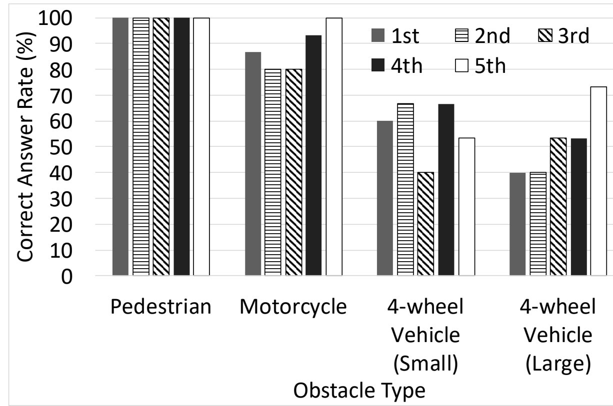


Figure 4. Correct answer rates for notification using haptic actuators.

B. Abilities of the Proposed Notification Systems

A study of our proposed conventional system proved it to be effective [8]. However, the vibrating motor used in our conventional system was unable to assign different vibration patterns to different obstacle types. Fig. 3 shows the correct answer rates for (a) the direction, (b) vibration intensity, and (c) both direction and intensity using the vibrating motor. When participants took longer than 5 s to respond, measured using a stopwatch, we considered it to be too slow and treated their answer as incorrect. In the correct answers shown in Fig. 3, the response times of all trials indicate that drivers could understand the information of the surrounding obstacles in less than 1 s. The correct answer rates for direction, intensity, and both direction and intensity in all route types were 84.4%, 72.6%, and 62.2%, respectively [8].

Fig. 3(a) shows that the drivers produced the highest number of correct answers when driving on the winding local road, followed by the arterials and the collector roads; however, as Fig. 3 (b) shows, the accuracy of the drivers' responses for the intensity were in the reverse order. This difference is likely because the vehicle's vibrations when travelling at low speeds could confuse the driver. The pressure between the vibrating motor and the buttocks could also change during the trials because the driver had to constantly control the accelerator and brake on the winding road. Nevertheless, the participants were able to determine the direction and intensity of over 50% of the vibrations when driving on the proposed seat.

For determining obstacle types, we applied the ACOUSTICHAPTIC™ actuator. Figs. 4 and 5 indicate the correct answer rates for the four types of obstacles. The graph also shows the rate for each trial by type. The vertical and horizontal axes in Figs. 4 and 5 show respectively the correct answer rate and vibration type of the waveforms shown in Fig. 2. Fig. 4 presents the correct answer rates based on the waveforms of the four obstacles types. As Fig. 4 shows, the correct answer rates improved during the trial, except for those for the four-wheeled vehicle (small). All participants were able to identify the pedestrian and motorcycle vibrations; however, they could only identify

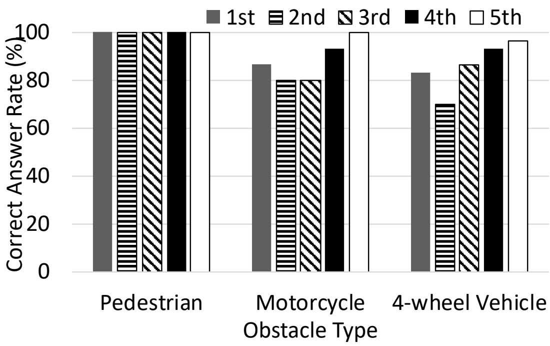


Figure 5. Correct answer rate for obstacle type notification on three obstacle types

50% of the other types of vibrations because the vibration patterns were too similar for them to sense the differences. However, after combining the large and small four-wheeled vehicles, the participants could detect the three patterns with high accuracy. Fig. 5 shows the results for three obstacle types, integrating the large and small four-wheeled vehicles. The correct answer rate reached over 90% at the fifth trial, as shown in Fig. 5.

IV. Modulation for Precise Notification

From the waveforms shown in Fig. 2, we generated modulated waves for precise notification. We determined the waves with a frequency an octave lower than the original wave as the modulated waves for more clearly feeling the differences of vibration. Figs. 6(a) and (b) show spectra of the original and modulated waves for large four-wheeled vehicles. The horizontal and vertical axes indicate the frequency and power spectrum, respectively. The spectra of the modulated waves consist of sine waves under 1 kHz.

TABLE II. SOUND VOLUME CORRESPONDING TO THREE STEPS OF VIBRATION INTENSITY.

	Volume (dB)
Small	−16
Medium	−8
Large	0

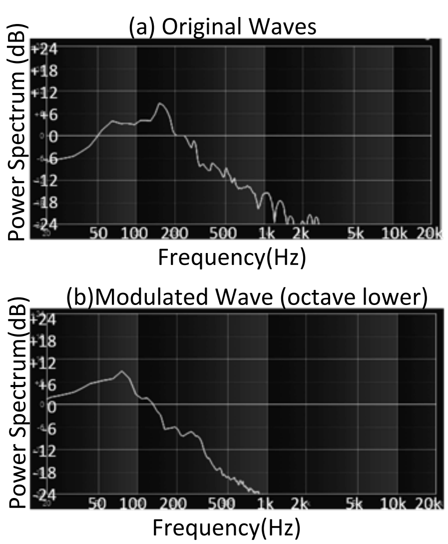


Figure 6. Sound spectra for 4-wheeled vehicles (a) before and (b) after modulation.

For the intensity expression, we utilized three intensity waves, as shown in Table II, which shows three vibrating volumes corresponding to three steps of intensity (*e.g.*, small, medium, and large). We applied 8 dB intervals between the three steps, and we prepared the signals with the three steps of volume on each modulated waveform for the three obstacle types: pedestrians, motorcycles, and four-wheel vehicles.

V. Experiments for Robustness Against Cushion Type

We evaluated the robustness against cushion type in a near-practical environment. As shown in Fig. 1, we mounted a vibrating car seat on a test vehicle for evaluation by five test drivers with considerable driving experience. The test drivers reported the vibration intensity, direction, and obstacle type when they sensed the vibration. Before the evaluation, the test drivers felt nine types of vibrations (*i.e.*, three intensities for the three obstacle types) at each actuator, shown as Actuator 1 to Actuator 5 in Fig. 1. The bold line in Fig. 7 indicates the experimental route, shown by Google Map [21]. The actuator is vibrated at random times while the test drivers drive on a circuit track, shown in Fig. 7, at speeds of less than 20 km/s. Answers were only considered valid when received within 5 s of the vibration.

A. Experimental Results

Fig. 8 shows the average ratings for comfort of each seat layer and material. The vertical and horizontal axes represent the average comfort rating and the cushion layer and material, respectively. The ratings were ranked according to the softness and resilience (high or low) of the seat cushion up to several layers. The test drivers tended to evaluate the seat based on whether they were conscious of the actuators.

The experimental results indicate the importance of retaining a high notification ability in a thick cushion even though a synthetic judgment is required for other evaluations,

Figure 7. Route for the driving experiment.

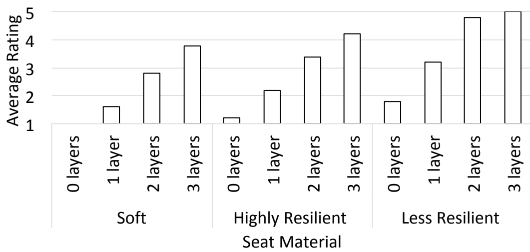


Figure 8. Comfort ratings for each seat layer and material.

such as ease of driving.

Figs. 9 to 12 show the results for robustness against cushion type and present the correct answer rates as relative values based on a correct answer rate for layer 0. The vertical and horizontal axes present the correct answer rate and cushion layer and material, respectively. Figs. 10 to 12 also present the standard deviations for all answers.

Fig. 9 shows the differences between the correct answer rates for the intensity, direction, and obstacle types between layer 0 and the other layers. The results show that the correct answers for the highly resilient cushion decrease as the layers increase; however, the other cushion materials maintain robustness even with an increased cushion thickness.

B. Intensity Expression

Fig. 10 shows the average differences in the correct answer rates calculated from the answers relating to the intensity of the vibrations (*e.g.*, small, medium, or large, as shown in Table II). The more layers the seat cushion has, the more the correct answer rates decrease, except for the less resilient cushion. A high standard deviation is obtained for the results of the less resilient cushion, although the trend of the correct answer rates for the layers is different. Therefore, it cannot be said that the number of correct answers will increase for more layers in the less resilient cushion.

C. Direction Expression

Fig. 11 shows the average differences in the correct answer rates, which are calculated from only the answers relating to direction (*e.g.*, left corresponding to Actuator 1 and right corresponding to Actuator 5, shown in Fig. 1). In this experiment, the test drivers gave the direction by stating "right", "right back", "back", "left back", or "left" when they noticed the vibration. The results shown in Fig. 11 confirm substantial differences between the different seat cushions.

D. Obstacle Type Expression

Fig. 12 shows the average differences in the correct answer rates calculated from the answers relating to the obstacle types (*e.g.*, pedestrians, motorcycles, or four-wheeled vehicles, shown in Fig. 5). The more layers there are, the more the correct answers decrease, except for the soft cushion. In the case of the soft type with 0 layers, it was difficult to judge when the actuators made direct contact with the buttocks based on the features of the waveform because the soft type of seat sank more easily than the other types. As a result, the correct answer rates on the soft cushion could be increased.

Based on the results shown in Figs. 9 to 12, the total correct answer rates are strongly influenced by the intensity of the vibrations; thus, the robustness is demonstrated without to the intensity steps of the vibrations.

VI. Hazard Notification System Mounted on a Real Vehicle

We discussed the accuracy of our system using Nissan MARCH as a real vehicle by experiments shown in Fig. 13 (a). Actuators are installed in a MARCH's seat, shown in Fig. 13 (b), based on conventional our experiments. Actuator 1 to 5, shown in Fig. 13 (a), are mounted in 40mm depth to touch the surface on the actuators at driver's buttock, and layout of each actuator is same as Fig. 1. Width between Actuator 1 and Actuator 5, Actuator 2 and Actuator 4 are 20cm,

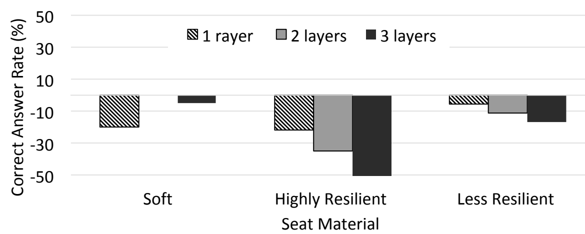


Figure 9. Correct answer rates for each seat material.

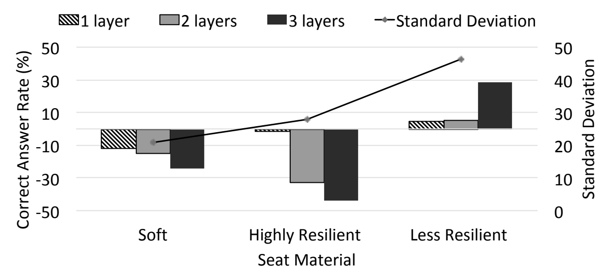


Figure 10. Correct answer rates for intensity for each seat material.

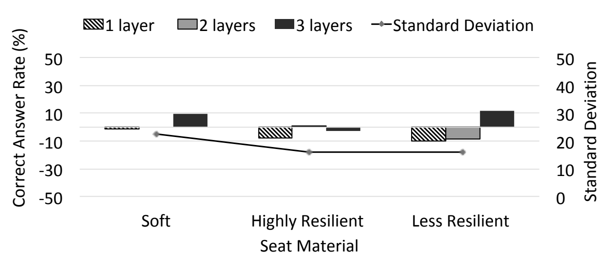


Figure 11. Correct answer rates for direction for each seat material.

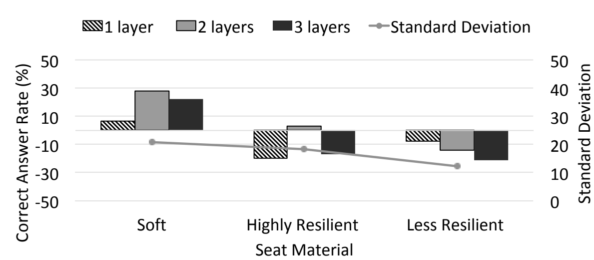


Figure 12. Correct answer rates for obstacle type for each seat material.

respectively, because standard buttock is mounted on the actuators. The seat is covered by normal MARCH's seat cover when we conduct experiments.

VII. Normalization by Using Deformed Waveform

To develop the correct answer rate, we considered sensitivity on each part of our buttock on each Actuator 1 to Actuator 5. Therefore, strength on each actuator was normalized by an experiment. Moreover, we changed to exaggerated waveform for more simply for easy to recognize.

A. *Deformed Waveform for Easy Understanding*

From the study written in Fig.5, higher correct answer rate on pedestrian was obtained than other obstacle types due to easier notification than other vibration waves. That is, we expect to develop accuracy on the obstacle types of Motorcycle and 4-wheel Vehicle if their waveforms are more different each other. Therefore, we utilize deformed waves for Motorcycle and 4-wheel Vehicle to develop the accuracy.

Fig. 14 (a) (b) shows deformed waves from Fig. 2 (b) and (d), respectively. Motorcycle is expressed by intermittent waveform as deformed motorcycle waveform, such as Fig. 14 (a), however interval is shorter than the waveform of the pedestrian as shown in Fig. 2 (a). In contrast, the waveform of 4-wheel vehicle is given by constant waveform to clearly express the deference between each other. The vibration waves are created by 55 Hz carrier waves of resonance frequency.

B. *Evaluation for Correct Answer Rate by Using Deformed Waves*

An experiment was conducted to evaluate accuracy of notification for obstacle types on moving vehicle. 9 drivers tested two types of vibration (*e.g.*, conventional waves shown in Fig.2 and our proposed deformed waves shown in Fig. 14) in this experiment. Actuators' strength was same as Table II. The participants drove the road shown in Fig.7 by the car shown in Fig. 13, and answer obstacle type from random vibrations within 3 obstacle types, 3 strength and 5 directions. Vibration timing is not told to drivers. Drivers answered when they feel the vibration. We treated a result as incorrect in case drivers did not answer within 5 seconds from start vibration.

Fig. 15 shows correct answer rate of conventional and the deformed waves on each obstacle type as the result of this experiment. A line in Fig. 15 shows the average of 3 obstacles. The correct answer rate was generally developed to over 90 % by using deformed waves. High notification ability for obstacle types was therefore confirmed by using proposed deformed waves from an experiment with real vehicle.

C. *Normalizing Method*

We evaluated accuracy on strength notification by using normalized strength. The strength was normalized based on conventional strength shown in Table II. The medium strength of -8dB shown in Table II is apply as medium decibel value between the large and the small decibel. We therefore also apply new medium strength of -5 dB, that is same magnification between large to medium and medium to small, and compare with conventional medium waves of -8dB.

We conduct the experiment for 9 drivers on the seat, shown in Fig.13, without driving. In this experiment, volume of the vibration waves for 4-wheel vehicles and Motorcycle are gradually decreased from conventional strength shown in Table II, after we gave the vibration on each strength for pedestrian by an actuator. Drivers answered the nearest strength of vibration for pedestrian from the decreased strength on other vibrations (*i.e.,* 4-wheel vehicles and Motorcycle).

D. *Experimental Result by Normalized Waves*

Fig. 16 shows an experimental result for strength on each obstacle. Horizontal and vertical axis denotes categories on each strength we gave for the drivers and average strength that drivers answered as relative output level from the strength for pedestrian, respectively. We also removed outlier of the answer, which is over from 1.5 times to interquartile range.

Fig. 16 obviously indicates the difference between the obstacle types especially 4-wheel vehicle against pedestrian.

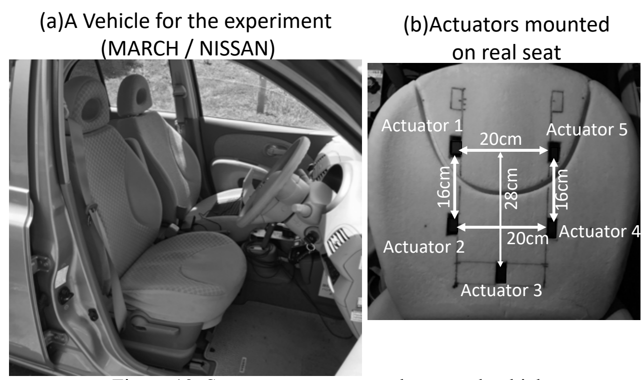


Figure 13. Seat actuators mounted on a real vehicle

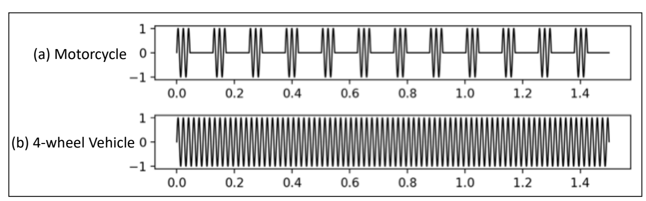


Figure 14. Deformed vibration waves for our system

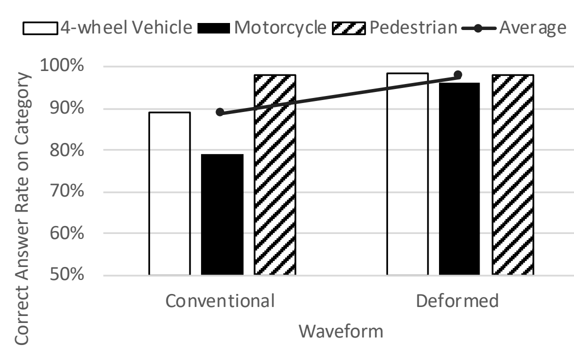


Figure 15. Accuracy on category notification by using deformed waves

By the experiment, -4dB and -3dB strength on 4-wheel vehicle and Motorcycle are obtained as same as the large strength for pedestrian, respectively. -12dB and -9dB strength on 4-wheel vehicle and Motorcycle are obtained as same as the medium strength (based on -8dB) for pedestrian, respectively. On the other hand, -9dB and -7dB strength on 4-wheel vehicle and Motorcycle are obtained in the case of the medium strength based on -5dB, respectively. Finally -19dB and -17dB strength on 4-wheel vehicle and Motorcycle are also obtained as same as the small strength for pedestrian, respectively. We applied these strengths as normalized strength.

E. Experiment for the Evaluation of the Normalization

We conducted an experiment for evaluating accuracy of notification by using the normalized strength. Others 9 drivers from previous experiments are participate with this experiment. The drivers answered which strength of large, medium, and small is vibrated from 45 pattern of vibration (*i.e.,* combination of 3 strength, 5 directions, and 3 obstacle types) at random

Experimental sequence is same as Fig 15. Fig. 17 shows correct answer rate of strength on each conventional setting (*i.e.,* strength is the same between all waveform of obstacle) shown in Table II, normalized setting based on decibel value, and normalized setting based on magnification, as the experimental result. Each rate of large, medium, and small strength in Fig. 17 is the total result of vibration with any direction and obstacle types. A line on Fig. 17 indicates average on 3 correct answer rates on the strengths.

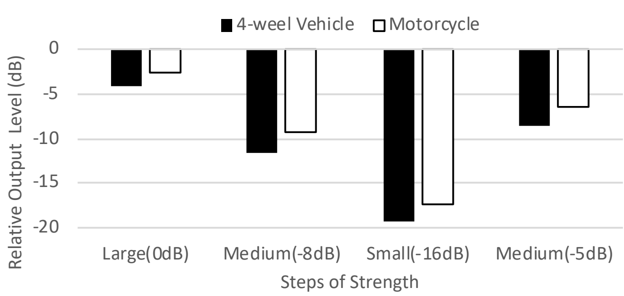


Figure 16. Sound volumes which research participants answer the strength as same as the strength of pedestrian

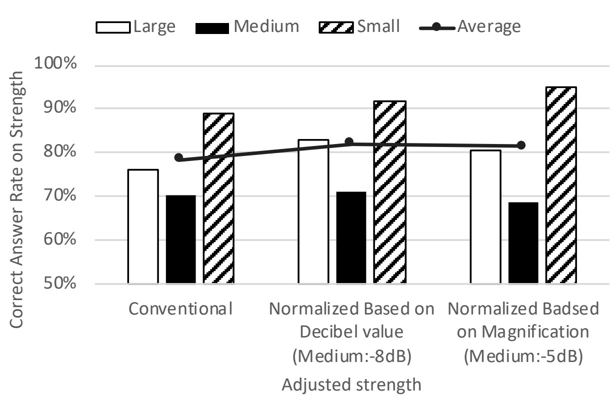


Figure 17. Correct answer rates on strength by normalized waves

Average of correct answer rate is under 80% in the case of conventional setting, however, by normalization, the average is increased. Focused on each strength, especially the strength of large and small is increased compared with medium strength. Correct answer rate of medium strength is smaller than other strength because driver answer incorrectly to both large and small. This tendency is almost same between any settings, and the correct answer rate is not improved. In the case of normalization based on magnification, the medium strength is nearer to large strength than the medium strength based on decibel value; therefore, the correct answer rate of large and small is decreased and increased respectively against normalization based on decibel value. We could not confirm a relation between correct answer rate and direction of actuators.

The notification with large strength is the most important to alert of all strength. We resulted normalization using medium decibel level is the most effective to realize high accuracy on notification because of the highest correct answer rate of large strength and average of all settings.

VIII. Combination with Conventional Visual and Sound Alert System

Although enough accurate alert system was realized, more accurate information for hazard will be expected in case of with conventional visual and sound alert system. We experimented for accuracy if conventional system is added to our proposed system.

A. Experimental Visual and Sound Notification System

Layout of experimental combination system is shown in Fig. 18. This system consists of PC for control alerts shown in Fig. 18(i), 3.5-inch monitor for indicating image shown in Fig. 18 (ii), a speaker for sound alert shown in Fig. 18 (iii), and our proposed system using ACHOSTIC HAPTIC Actuators shown in Fig. 18 (iv). Information of obstacle types, and distance and direction from the obstacle is sent from the PC to these media.

We utilized 4-wheel vehicle, motorcycle, and pedestrian sign as shown in Fig. 18 (a), (b) and (c), respectively. 3 dots

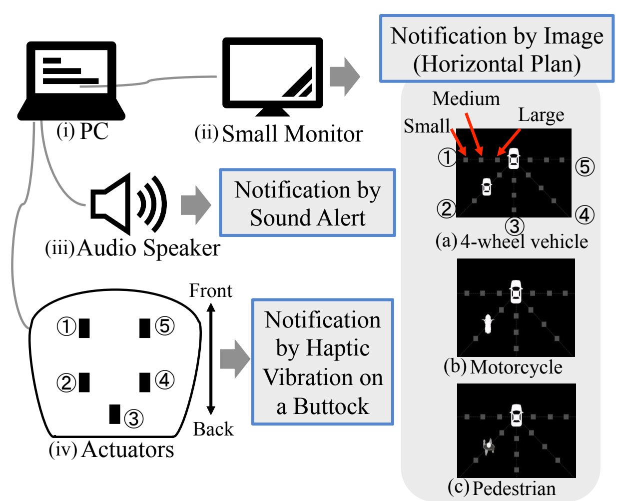


Figure 18. System layout for the combination system

corresponding strength are displayed on each direction depending on Actuator 1 to Actuator 5, shown as 1 to 5 in Fig. 18. Vibration waveform is the same as normalized waves based on decibel shown in Fig. 16. Sound alert is utilized as same as conventional practical systems.

B. Evaluation for Combination Notification

We experimented to evaluate effectiveness of the combination system by based on drivers' subjective feelings. 8 drivers, whose age is 18 to 22, ware participated with this experiment. The system alerts information simultaneously by 3 medias when they drive a paved route shown in Fig. 7. Drivers answered not only strength, obstacle types, and directions, but also a questionnaire about priority between the alert medias to decision, and interference and necessity of our notification method using vibration.

As a result, by using multimedia notification, correct answer rates of all strength, direction, and obstacle types increased near to 100% even if low correct answer rate was obtained by notification with only vibration. All driver answered when we gave alerts in this experiment.

Fig. 19 shows order in which drivers noticed from 3 medias. Horizontal and vertical axis shows a kind of medias on each situation and the number of drivers, respectively. The situation in the Fig. 19 (a) and (b) is normal driving situation with less risk for incidents and situation when a driver pay attention to other obstacles, signal, sign, or road condition, respectively. Most drivers ware firstly notified by sound and vibration in the case of normal and attention situation as shown in Fig. 19 (a) and Fig. 19 (b), respectively. In the case of Fig. 19 (b), drivers finally confirmed the alert by display because driver's payed attention to others by eyes.

Fig. 20 also shows the driver's most decisive media between 3 medias on each first awareness, direction, strength, and obstacle types. Horizontal and vertical axis shows a kind of medias on each situation and the number of drivers, respectively. Most drivers firstly awared by sound of any situation. In the case of Fig. 20 (a), strength, direction and obstacle types are decided by image shown in display, however, in the case of Fig. 20 (b), drivers chosed as decisive media by vibration more than Fig. 20 (a). Driver is difficult to watch a display when drivers pay attention around own vehicle. In this experiment, drivers recomfirmed by display on normal situation shifted from the attention situation. Our proposed system using vibration can imediatelly alert the obstacle type, distance and direction from obstacle, therefore the "Vibration" shown in Fig. 20 (b) was increased. Therefore, multimedia notification including our proposed method is possibilly effective for emergent notification with an insident especially in the case of paying attention to others from this experiment.

Fig. 21 and Fig. 22 indicate answer for a questionnaire by drivers. Fig. 21 shows subjective impressions for distinguish between notifications of strength, direction, and obstacle type. In the case of Fig.18 (a), 87% of driver can be notice with extremely distinguished by visual image. However, Fig.18 (b) shows 1 step-decreased result of distinguishes from Fig.18 (a). Fig. 22 also shows the result of questionnaire about recognition and interference by using multimedia. As shown in Fig. 22, 63 % of drivers answer that it is easier to recognised information by both vibration, visual images and sound alert than alert with only sound and visual images. 88 % of drivers answer that they can drive without interference even if both visual, sound and vibration is activated. Therefore, notification by using vibration is

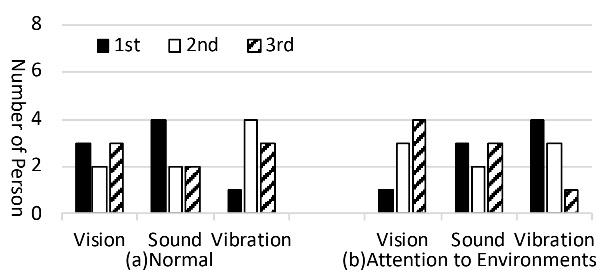


Figure 19. Priority of notifying media for notification

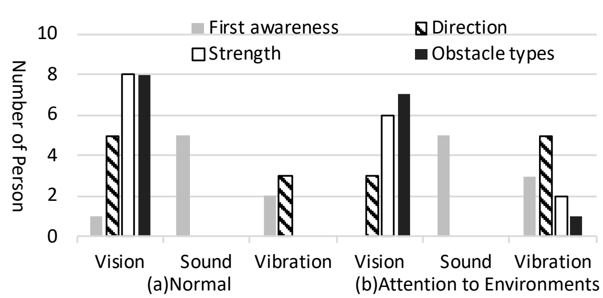


Figure 20. The most decisive media

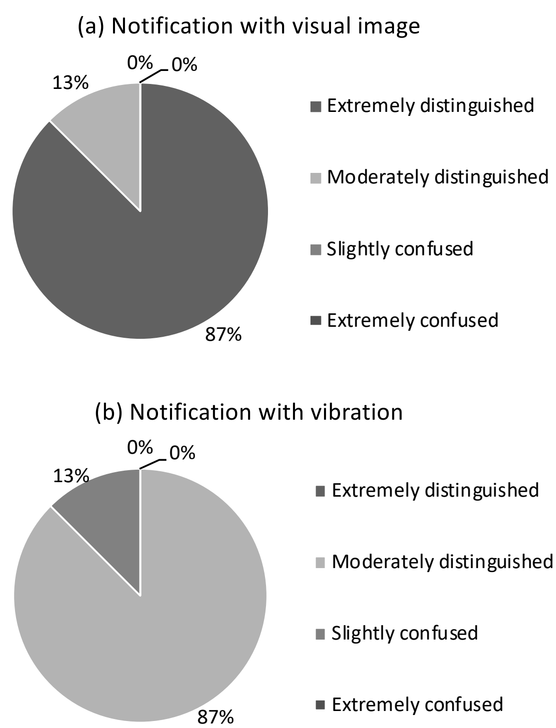


Figure 21. Subjective impression about distinguish for (a) visual image and (b) vibration

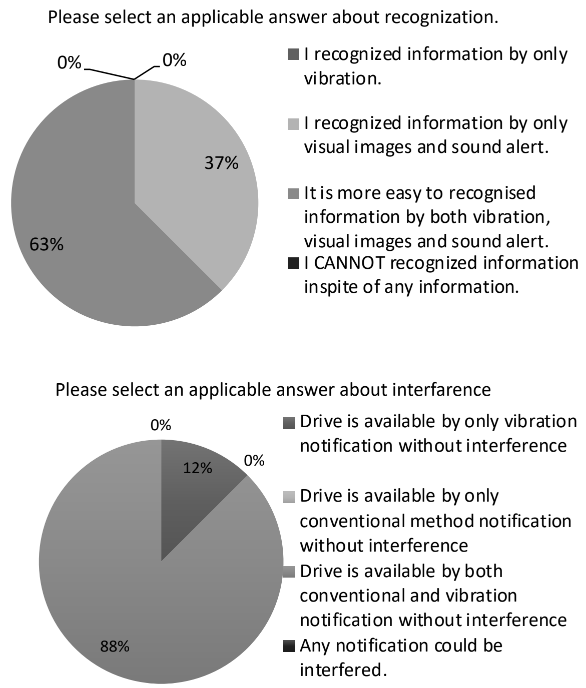


Figure 22. Questionnaire for needs and interference

effective in the case of multimedia system.

IX. Conclusion

We examined the robustness of a vibro-tactile device by collaborating with a car manufacturer to install acoustic haptic actuators into the seat cushion of an actual automotive vehicle.

We proposed a vibro-tactile notification system using vibrating motors to notify drivers of hazards around a vehicle, which were sensed using conventional sensors. The effectiveness of this method was evaluated from the viewpoint of resolution of intensity and direction and robustness against cushion type for determining the road conditions. The vibration pattern also enabled drivers to recognize the type of hazard, such as an approaching pedestrian or motorcycle.

We conducted several experiments involving driving on public roads in a car with seven vibrating motors installed under the driver's seat. By applying acoustic haptic actuators as a vibro-tactile device, test drivers could detect three types of vibrating patterns, indicating different types of obstacles: pedestrians, motorbikes, and four-wheeled vehicles. We also determined that drivers' notification of the intensity, direction, and hazard type could be improved over time because they could learn from experiencing the vibration alerts.

The results indicated the high potential of using a haptic sensation device to notify drivers of obstacles in their blind spots by creating a vibration against the buttocks. The experimental results, shown in Figs. 9 to 12, illuminated that the intensity of the vibration, which should indicate the level of the hazard, could not be considered in the robustness test. By reconfiguring the intensity, as shown in Table II, the robustness could be improved, which we will investigate in future work.

We also evaluated our system by using real vehicle. We conduct experiments to evaluate for proposed deformed vibration waves and normalizing method between the waveform. From this experiment, an accuracy of our notification method is developed to over 80% by using normalized waves based on decibel value.

In the case of applying for multimedia system with sound speaker and visual display, more accuracy of information for obstacle around own vehicle was obtained. Moreover, interference by using multimedia was not confirmed and multimedia notification including our proposed method is possibilly effective for emergent notification with an insident especially in the case of paying attention to others from this experiment.

The proposed system is expected to reduce accidents by notifying drivers of other drivers and obstacles. In future works, we will conduct these experiments using more test drivers to compare elderly people with very young people or to observe its effectiveness with truck drivers and people with different levels of attention or tiredness. We will also evaluate the operational difficulties of the system in case of an emergency. We will also corroborate the visual and audio notifications of our system and examine the effectiveness for making quick driving decisions with intuitive notifications.

Acknowledgments

This work was supported by MEXT KAKENHI Grant Number JP 16723884 and Foster Electric Company, Limited.

References

[1] A. Suzuki, Y. Murata and S. Fujimura, "Robustness Against Hazard Notifications Around a Vehicle Using Seat Actuators," in The Eighth International Conference on Advances in Vehicular Systems, Technologies and Applications (VEHICULAR 2019), Rome, Italy, 2019.

[2] A. Hevelke and J. Nida-Rümelin, "Responsibility for crashes of autonomous vehicles: an ethical analysis," Science and Engineering Ethics., vol. 21, no. 3, pp. 619–630, June 2015.

[3] National Public Safety Commission and National Police Agency, "Traffic accidents situation (2017Oct.)," Dec. 2017. [Online]. Available from: http://www.e-stat.go.jp/SG1/estat/GL32020101.do?method=xlsDownload&fileId=000008051596&releaseCount=1 2019.06.24

[4] Subaru of America, Inc., Dec. 2017. [Online]. Available from: https://www.subaru.com/engineering/eyesight.html 2019.06.24

[5] Toyota Motor Sales, U.S.A., Inc., "SUZUKI GSX-S1000 normal idling sound motorcycle," Dec. 2017. [Online]. Available from: https://www.toyota.com/safety-sense/animation/pcs 2019.06.24

[6] Infineon Technologies AG, Dec. 2017. [Online]. Available

from: https://www.infineon.com/cms/en/about-infineon/press/market-news/2009/INFATV200905-058.html?redirId=46231 2019.06.24

[7] I. Rock and J. Victor, “Vision and touch: an experimentally created conflict between the two senses,” Science, vol. 143, no. 3606, pp. 594–596, Feb. 1964.

[8] A. Suzuki, Y. Murata, and M. Hayashi, “Notification of hazards around a vehicle using seat actuators,” in 25th ITS World Congress, Copenhagen, Denmark, 17–21 Sep. 2018.

[9] L. Akonso, V. Milanés, C. Torre-Ferrero, J. Goboy, J. P. Oria, and T. de Pedro, “Ultrasonic sensors in urban traffic driving-aid systems,” Sensors, vol. 11, pp. 661–673, Sep. 2011.

[10] J. K. Suhr and H. G. Jung, Fully-automatic Notification of Various Parking Slot Markings in Around View Monitor (AVM) Image Sequences, 15th International IEEE Conference on Intelligent Transportation Systems (ITSC), Anchorage, USA, 1244-1299, Sep. 2012.

[11] L. Qingquan, C. Long, L. Ming, S. Shih-Lung, and N. Andreas, “A sensor-fusion drivable-region and lane-detection system for autonomous vehicle navigation in challenging road scenarios,” IEEE Vehicular Technology Society, vol. 63, no. 2, pp. 540–555, Feb. 2014.

[12] M. Tada, S. Sekiguchi, T. Nisimatsu, and E. Toba, “Measurement and evaluation of sitting comfort of automotive seat,” in Instrumentation and Measurement Technology Conference, 1998. IMTC/98. Conference Proceedings. IEEE, St. Paul, MN, USA, USA, 1998.

[13] A. Cassinelli, C. Reynolds, and M. Ishikawa, “Augmenting spatial awareness with the haptic radar,” i10th IEEE International Symposium on n Wearable Computers, Montreux, Switzerland, 2006.

[14] R. W. Lindeman, Y. Yanagida, J. L. Sibert, and R. Lavine, “Effective vibrotactile cueing in a visual search task,” IFIP TC13 International Conference on Human–Computer Interaction (INTERACT '03:), Zurich, Switzerland, 2003.

[15] C. C. Ltd., “Vehicle notification device, vehicle notification method and notification signal”. Japan Patent WO2017090355 A1, 1 6 2017.

[16] Alps Electric Co., Ltd., Dec, 2017. [Online]. Available from: http://www.alps.com/prod/info/E/HTML/Haptic/ 2019.6.24

[17] M.-studio. official Channel, “Recommend sound effect, Footstep-Leather shoes 04 On-Jin” [Online]. Available from: https://www.youtube.com/watch?v=A5fGVbsHOfI 2019.06.25.

[18] GYANTARO, “Vellfire V6 3.5L 0-138km/h Full Acceleration Test (Alphard 30 series),” [Online]. Available from: https://www.youtube.com/watch?v=WI4voS7YF6Y 2019.06.25.

[19] TAKAPON CHANNEL, “Saihi tourist bus JAPAN” [Online]. Available from: https://www.youtube.com/watch?v=EsJnG-E22ss 2019.06.25.

[20] ahotakuro, “SUZUKI GSX-S1000 Normal Idling Sound Motorbike” [Online]. Available from: https://www.youtube.com/watch?v=c6pK2z9ANRk 2019.06.25.

[21] Google, [Online]. Available: https://www.google.com/maps/@39.7560755,141.1615498,13.42z 2019.06.25.

Driver Emotional States & Trust: Interactions with Partial Automation On-Road

Liza Dixon
Hochschule Rhein-Waal
University of Applied Sciences
Kamp-Lintfort, Germany
email: lizadixon@gmail.com

William M. Megill
Hochschule Rhein-Waal
University of Applied Sciences
Kleve, Germany
email: william.megill@hochschule-rhein-waal.de

Karsten Nebe
Hochschule Rhein-Waal
University of Applied Sciences
Kamp-Lintfort, Germany
email: karsten.nebe@hochschule-rhein-waal.de

***Abstract*—Many vehicles on-road today are equipped with Advanced Driver Assistance Systems (ADAS) which enable a driver to handover primary driving tasks to the vehicle under specific conditions, provided the driver continues to supervise the system (Level 2 automation). Various tools and methods are used in the study of human-machine interaction with vehicle automation in order to assess a driver's experience and interactions with a system. When used in-vehicle, Facial Emotion Recognition (FER) offers researchers the possibility of a quantitative reading of the driver's changing emotional state in response to interactions with the system. This paper presents a method of correlating FER data post-drive with participants' reported feelings of trust in the system. FER visualizations of the duration of the test drive sessions as well as visualizations of specific driving events are presented. Challenges in the use of FER in-vehicle, "in the wild" (on-road) are also discussed. Participants with a gain in trust post-drive and those with a loss in trust post-drive more frequently displayed the emotions *happy* and *angry*, respectively. Results indicate that trust increases after a user's first experience with an ADAS and further that FER may be predictive of user trust in automation.**

Keywords—Facial Emotion Recognition; Driver Emotions; Advanced Driver Assistance Systems (ADAS); Human-Machine Interaction; Automation.

I. INTRODUCTION

Exponential improvements in computing speeds, computer vision, and machine learning over the past decade are fundamentally changing what it means to "drive a car". Inside the vehicle, a revolution is taking place—one in which the primary role of the driver is shifting to that of a passenger [1]. This shift in the role of the user presents substantial challenges in the acceptance of this technology and raises important social, ethical, and legal concerns about the future of road transportation.

From a road safety perspective, the incentive for user acceptance of vehicle automation is clear—a majority of auto accidents are due to human error, killing 1.35 million people each year and leaving up to 50 million injured or disabled, internationally [2]. In addition to a humanitarian concern, acceptance is also an economic concern for companies heavily invested in the research and development of vehicle autonomy [3]. Failure to support users in their exchange with Advanced Driver Assistance Systems (ADAS) on-road today, "will become increasingly costly and catastrophic [4]," as vehicle automation grows in its capabilities and prevalence.

The term Advanced Driver Assistance System refers specifically to the current state of the art in production vehicles, also known as a Level 2 (assisted or partially automated) system, whereas Automated Driving System (ADS) refers to vehicles that are conditionally, highly or fully automated (Level 3, 4 or 5, respectively). The pervasiveness of partial automation (Level 2) in the form of ADAS and the dawn of conditional automation (Level 3) in production vehicles necessitates a robust understanding of Human-Machine Interaction (HMI) challenges in vehicle automation [5].

Each level of vehicle autonomy requires differing levels of supervision from the user and presents various challenges from a HMI perspective [6]. While an ADAS supports the driver in longitudinal and lateral control of the system under specific conditions, the driver is still required to remain focused on the road and prepared to take control of the vehicle at all times. Because these systems are characterized by their limitations as *assistive* systems [5], ADAS involves frequent handovers of control between the system and the user, leaving substantial room for error [6], [7].

Studies have been conducted to gather insight regarding user sentiment towards vehicle autonomy, the results of which point to trust as a major factor in the acceptance of the technology [8]–[10]. Trust is defined by Mayer et al. [11] as an attitude; it is not risk-taking, "but rather it is a willingness to take risk." Adjusted for the context of automated systems by Körber, trust is "the attitude of a user to be willing to be vulnerable to the actions of an automated system based on the expectation that it will perform a particular action important to the user, irrespective of the ability to monitor or to intervene [12]." Trust affects reliance on automation, and reliance aids the user in navigating the complexities of automated systems, especially when the context of use demands adaptive behavior, as is the case with ADAS [4]. Furthermore, a study of trust in automation by Miramontes et al. [13] concluded that people with "high emotional stability…reported higher levels of trust in automation."

Whether it is getting cut off in rush-hour traffic or a series of traffic lights ahead signaling green—the act of driving can be an emotional roller-coaster. In turn, a driver's emotional state influences their driving performance. A literature review of emotion on the road by Eyben et al. [14] reveals that "happy drivers are better drivers" and that "aggressiveness and anger are emotional states that extremely influence driving behavior and increase the risk of causing an accident." As a tool, Facial Emotion Recognition (FER) can facilitate a deeper understanding of the user experience in-vehicle [15], [16] and aid in the identification of drivers with a tendency towards *happiness* and those with a tendency towards *anger* on the road. In theory, it is possible that FER may be indicative of a users' trust in automation.

While closed courses and simulators provide a stable research environment for human-machine interaction with vehicle automation, they are not fully aligned with the context of use, nor the state of the art. The release of one's personal safety to the system occurs exclusively while engaging with partial automation in an on-road setting. Hence, existing research is unable to specify how exposure to and experience with a Level 2 system on-road might impact user trust. While Facial Emotion Recognition is typically used to assess the user experience, it has yet to be applied in an on-road study in correlation with driver trust in vehicle automation. In order to address this, the following research questions were explored in an experiment:

Q1: How does a driver's first experience with an Advanced Driver Assistance System on-road affect their level of trust in the system?

Q2: Which emotions (*neutral, happy, surprise, angry*) do first time drivers of an ADAS display and is there a relationship between the emotions displayed and their reported levels of trust in the system?

Q3: Which emotions (*neutral, happy, surprise, angry*) accompany specific assisted driving events?

Incorporating a mixed-method approach utilizing verbal trust scores, the Trust in Automation questionnaire [12], and Facial Emotion Recognition, and qualitative/observational data, it was expected that the results of the experiment would indicate the following:

H1: Participants will report higher levels of trust in ADAS after their first experiential drive with an ADAS.

H2: FER analysis will reveal a relationship between a participant's Trust in Automation score and their emotions displayed during the drive.

H3: Participants will display varying emotions according to the driving scenario and behavior of the ADAS, including an initial emotional response to a driving event, followed by emotional resolution i.e., a return to their respective normal emotional state as defined by FER.

Section II of this paper provides related works and background information, and Section III discusses participant demographics, the technical capabilities of the vehicle utilized for test drive sessions, and the experiment procedure. Results are reported in Section IV and Section V summarizing qualitative and quantitative findings, respectively. Section VI is a discussion of the results, followed by Section VII, which outlines the limitations of this study. The paper concludes with Section VIII, which offers an outlook and future work.

This paper is an extension of work originally presented in VEHICULAR 2019: The Eighth International Conference on Advances in Vehicular Systems, Technologies, and Applications [1]. This work varies from the original in the following ways, adding: 1) more background on driver emotions 2) an additional research question (Q3) 3) more detailed information about the participants in the study 4) the results of an exploratory analysis of each participant's emotional journey throughout the experiment (Section V.C) and during a selection of specific driving events (Section V.D.) 5) discussion about the limitations and difficulties of working with FER in-vehicle (Section VI and Section VII).

II. Related Work

Experiments to measure trust in vehicle automation have been carried out in both closed courses [11][12] and simulators [19], [20]. For example, in an experiment with 72 participants, Gold et al. [19] utilized a driving simulator modeling a Level 3 system to "investigate how the experience of automated driving will change trust in automation and the attitude of the driver towards automation." A questionnaire was administered before and after a 15-20 minute driving experience. Gaze behavior was also recorded in an effort "to measure a change of trust by a change in [eye] scanning behavior." The results of this study revealed that participants reported a higher level of trust in automation after the driving experience, however gaze behavior could not be established as a valid measurement.

In an experiment using a Wizard of Oz setup (simulating an automated vehicle), Ekman et al. [18] explored a mixed-methods approach for the assessment of trust during a 15 minute drive on a closed course with 18 participants. The results of the study indicated that "data should not only be collected at the very end of a trial only but be complemented with data collection also during a trial, in particular in relation to events that may influence and contribute to a user's overall experience."

Researchers have developed frameworks, models and scales for the assessment of user trust in vehicle autonomy. Ekman et al. [17] constructed the Lifecycle of Trust (LCoT) framework, to serve as a tool for HMI design. The LCoT identifies 11 trust-affecting factors throughout the *Pre-Use Phase* (*Implicit/Explicit Information*), *Learning Phase* (all activities from *Entering the Vehicle* to transitions from *Manual to Automated Control*, to *Exiting the Vehicle*) and *Performance Phase* (covering *Continuous Usage, Change of Context* & *Incidents*). Validation of LCoT factors, specifically through the *Pre-Use* and *Learning Phases* are an area of interest for this study, as it is the most current, comprehensive framework for understanding the development of trust in automation.

Based on empirical research, Jian et al. [21] developed the "Checklist for Trust between People and Automation" a 7-

point Likert scale comprised of 12 questions designed for use as a general scale in any area where human-automation interaction occurs. Based on this, the work of Mayer et al. [11], Lee & See [4] and others, Körber [12] developed a refined model of Trust in Automation (TiA) with an accompanying 19-item, 5-point Likert scale questionnaire covering the following factors: *Reliability/Competence, Understanding/Predictability, Intention of the Developers, Familiarity, Propensity to Trust* and *Trust in Automation*. The questionnaire features questions, such as "The system is capable of interpreting situations clearly," (*Reliability/Competence*) and inverse items such as "The system reacts unpredictably," (*Understanding/Predictability*) which correspond to the underlying factors. To the knowledge of the authors, this questionnaire has yet to be applied in a study of trust in partial automation on-road.

A. Driver Emotions

Facial Emotion Recognition via the analysis of facial expressions extracted from images/video frames is of growing interest to HMI researchers in the area of automated driving. FER consists of three main events: 1) face and facial component detection, 2) feature extraction, and 3) expression classification [22]. The facial expressions which are associated with the emotions *happy* (*joy*), *anger*, and *surprise* are thought to be the most relevant in the context of automated driving and are used by commercial software companies in their analysis [16]. Studies have confirmed that the emotions *happy* and *angry* are the most influential on how a car is driven [23].

When used in-vehicle in real time, systems can perhaps use FER data to align its behavior with the changing emotional state of the user. This concept of system adaptation, was originally presented by Picard [24] and is known as *affective computing*. Picard suggests that computers which "...sense and respond to users' emotional states may greatly improve human-computer interaction" [4]. Affective computing by means of FER is of particular interest in vehicle automation, as improving the naturalness of interaction with the system facilitating trust and acceptance [25]. Further, according to Lee & See [4], understanding "Emotional response to technology is not only important for acceptance, it can also make a fundamental contribution to safety and performance."

A well-designed, supportive system assists the driver in achieving their goals by taking over the monotonous tasks of driving and correcting errors such as inattentiveness, smoothing out the driving experience. However, a system which is not well-aligned with the user or behaves unexpectedly may cause friction, resulting in displeasing emotions and disuse of the system. Acceptance of automated driving systems and appropriate reliance from the user is crucial to ensure its efficient and safe use [4]. Understanding the emotional states of a driver before and after driving events, while operating an ADAS on-road may support affective computing techniques in next generation vehicles [14] and in turn, trust in automation.

When used as part of a mixed-method approach in post-production, FER may therefore enable the observation of correlations between driver emotional states, vehicle behaviors and reported trust in automation.

III. Methods

A total of n=10 participants were introduced to the same Level 2 vehicle and completed one individual test drive session. All participants completed their test session within the same two-week period. The driving route included driving time on the autobahn (including a construction zone), country roads and in urban settings. Each experiment session lasted 1 hour and 30 minutes, approximately an hour of which was driving time. The same moderator accompanied all of the participants; participants were not explicitly told to activate the ADAS. A pilot test was conducted to refine the experiment structure and equipment, after which it was determined the route did not include enough autobahn time and was therefore revised.

A. Participants

Of the ten participants selected for this study, there were six females and four males. All participants were members of the university community and were recruited via email and flyer. Participants were screened prior to the experiment session to ensure they met specific requirements for the study: holding a valid driver's license, experience with automatic transmission, have no prior experience with the vehicle class (Mercedes-Benz GLC), not own or regularly operate a Mercedes-Benz, have no prior first-hand experience with ADAS, any semi-autonomous or autonomous vehicle systems (including for example: autopilot systems, adaptive cruise control or lane keeping assistants. Excluding: standard cruise control/speed limiters, back up cameras or blind spot assistants).

The driving route was designed as a loop, beginning and ending at each participant's respective campus. All participants drove the same section of the autobahn, with additional driving time in country road and urban settings which varied slightly based on the participant's starting location. Participant 1 (going forward, participants are referred to as "PX" where X is their individual coded reference number), P2 and P3 began at campus A, while P4-P10 began at campus B.

TABLE I. PARTICIPANT PROFILE

Participant	Age	Gender	Education	*Technical
P1	>30	F	Vocational	No
P2	<30	F	Vocational	No
P3	<30	F	Vocational	No
P4	>30	M	Masters	No
P5	>30	F	PhD	Yes
P6	>30	F	Masters	Yes
P7	<30	M	Vocational	No
P8	<30	M	Vocational	Yes
P9	>30	M	Masters	No
P10	>30	F	Masters	Yes

*Denotes whether or not the participant had a computer science or other engineering background.

The mean age of the participants was M = 31.66 years (SD = 9.17, ranging from 20 to 48 years old). Six of the participants had been driving for over ten years while four had been driving for ten years or less. Educational level was split 50/50 between the participants, half holding a master's degree or above and the other half having received vocational training. Table I profiles each participant.

1) Disclosure: Participants were informed only that they would be taking part in a study of Human-Machine Interaction in ADAS which involved an on-road test drive. The focus of the study being specifically about their trust in the ADAS was intentionally withheld from the participants. It was not disclosed explicitly nor by accident (e.g., titles removed from trust questionnaire, discussion about experiment sessions prohibited) in order to mitigate the *Hawthorne Effect*. This effect refers to the inclination of research participants to adjust their behavior and act in a way that they believe is aligned with the expectation of the moderator [26]. This decision was also made in part due to the high cognitive demands of the experiment [14] (driving an unknown vehicle with unfamiliar technology on public roadways, while under observation) and to obtain unbiased and natural reactions in any participant commentary related to the discussion of trust in the system.

B. Vehicle

The same Mercedes-Benz GLC-250 4Matic was driven by all participants. This vehicle was equipped with the Driving Assistance Package Plus option which includes an Advanced Driver Assistance System (sub-systems relevant to this study are listed in Table II). These features qualify the vehicle as a partially automated, Level 2 system [5].

TABLE II. SELECT DRIVER ASSISTANCE PACKAGE PLUS FEATURES

Feature	Function	Active
Distance-Pilot DISTRONIC with Steering Assist and Stop&Go Pilot	"Autonomous intelligent cruise control system" able to accelerate and decelerate according to traffic conditions. Steering interventions help the driver stay in lane. The system can follow the vehicle ahead even where there are no or unclear lane markings (<130 km/h).	0-200 km/h, driver activated
Hands-Off Warning	A haptic (steering wheel vibration) and graphic warning (in the multi-function display, next to the speedometer), alerts the driver to return their hands to the wheel. If this is not heeded, it is enhanced via an auditory warning tone.	Active with DISTRONIC
Active Lane Keeping-Assist	Detects unintentional lane drift by monitoring road markers. Can tell if the vehicle veers out of lane without signaling, and will vibrate the steering wheel. Brakes individual wheels for correction, keeping the vehicle within the road markers.	60-200 km/h, (conditional)
PRE-SAFE® Brake with Pedestrian Detection	Able to detect pedestrians ahead and will apply the brakes automatically.	Up to 50 km/h
Traffic Sign Assist	Identifies traffic signs and speed limits on the instrument display via camera and GPS data.	Always active.

Source:[27].

The purpose of this study is not to cross-compare various technologies, but rather, to analyze the inherent trust in a particular vehicle's systems, holding this as a constant.

C. Procedure

In order to ensure consistency and objectivity between the experiment sessions, the moderator adhered to a set procedure (see Table III) and script. At the start of the session, the moderator greeted the participant outside of the vehicle in the parking lot. This is when the participant was first exposed to the make and model of the vehicle. The participant was invited to enter the vehicle, where they were then interviewed regarding their initial impressions of the vehicle, Mercedes-Benz, thoughts about ADAS, vehicle autonomy, and their expectations of the system, including their initial feelings of trust in the system. They were then asked to give a verbal rating of trust in the ADAS on a scale from 1 to 5 (1=low, 5=high). Next, they watched an introductory video featuring original content from Mercedes-Benz, which was produced for this experiment by the researchers to reflect the capabilities of the specific vehicle used for testing. The participant was informed that they were in full control of the vehicle at all times and responsible for obeying all traffic laws and posted signs. Next, the Trust in Automation questionnaire (modified from [12]) was administered to the participant in their native language (German or English, translation from [12]). After, they were encouraged to ask questions, to ensure their understanding of the system's functions and capabilities. They were asked to rehearse how to activate/deactivate the system while the car was parked. Following the introduction, each participant was asked for a second time to give a verbal rating of trust in the ADAS. Trust in Automation, Pre-drive vs Post-drive.

TABLE III. EXPERIMENT PROCEDURE

Pre-Drive		Drive	Post-Drive
Introduction I	***Introduction II***	***Test Drive***	***Closing***
1) Interview 2) VTS #1	1) Intro video 2) TiA #1 3) Interview 4) VTS #2	1) Planned route 2) Think aloud 3) FER	1) TiA #2 2) Interview 3) VTS #3

VTS = Verbal Trust Score, TiA = Körber's Questionnaire for Trust in Automation, FER = Facial Emotion Recognition

As the participants began the test drive with the route pre-programmed into the vehicle's GPS, the GoPro cameras were activated at 60fps. The driver-facing camera was mounted to the windshield to the right of the steering wheel for later FER analysis. The driving scene (roadway ahead), multi-function display, and participant's interaction with the system's interface was were captured by a second camera mounted behind/next to the driver's right shoulder.

During the test drive, the moderator did not give any tasks to the participants other than to follow the route on the GPS. The moderator played an observatory role, giving instruction only when prompted (e.g., clarifying a system limitation). Participants activated the system only as they felt comfortable, in the appropriate conditions and were encouraged to think aloud [4] while doing so. Participants were asked to state aloud whether they or the car was performing certain actions (steering, braking, acceleration/deceleration) throughout the drive and to share their thoughts on the vehicle's behavior as it occurred. Top speed with ADAS active was recorded for each participant as well as adjustments in posture (positioning of hands, arms and feet on/off pedals). Immediately following the drive, the TiA questionnaire was administered a second time. Participants then completed a post-drive interview and gave a final verbal rating of trust in the ADAS on a scale from 1 to 5 (1=low, 5=high), based on their experience. All interview audio was recorded for later reference.

1) Data Analysis: The responses to the TiA questionnaire were scored following the procedure used by the System Usability Scale [28]. Adjusted for the number of questions, responses were reverse coded, added together and then multiplied by a factor to convert the original scores of 0-68 to a 0-100 value, in order to better identify discrepancies in participant's pre-use and post-use scores (the factor *Familiarity* was removed from analysis, as all participants were selected purposefully to have no prior experience with the technology). A Wilcoxon Signed-Rank Test was used to examine differences in pre-drive and post-drive, reverse coded TiA questionnaire medians. This method was chosen as it is appropriate for the comparison of medians in ordinal data from related groups with a symmetrical distribution [29]. Wilcoxon was performed for all TiA factors together (*Reliability/Competence*, *Understanding/ Predictability*, *Intention of the Developers*, *Propensity to Trust* and *Trust in Automation*) and for each factor's respective set of questions. Friedman's Test (adjusted for ties) was used at to analyze shifts verbal trust scores (pre-introduction, post-introduction and post-drive). Friedman's was selected as the data is ordinal, came from a single group measured at three intervals, and there are no interacting effects between the groups [30], [31]. Statistical analysis and plotting of TiA and verbal trust scores was completed in RStudio [32] using the *stats* [33], *agricolae* [34], and *ggplot2* [35] packages.

2) Facial Emotion Recognition: Driver facing video footage captured at 60fps was processed by a convolutional neural network (CNN) with 3 convolutional layers and two fully connected layers (including the output layer). The CNN was trained for the facial emotion recognition of seven emotions [36], however a reduced set of emotions was selected for analysis: *neutral*, *happy*, *surprise*, and *angry* [14], [16]. Classification performance using this set of emotions was reported 81% accurate by Mathworks MATLAB 2018b [37], which was used to run the network and output the data in text files. The text files were then compiled, cleaned and analyzed in RStudio.

The analysis returned a value from 0-1.0 for each emotion, (where *neutral*, *happy*, *surprise* and *angry* share a portion of the 1.0 value) for each frame and output the data in text files. The text files were then compiled and plotted in MATLAB using movmean [38]. FER scores were reviewed for each test drive as a whole, as well as for specific driving events. These events included: removing hands from the wheel (Hands Off), strong deceleration and steering through curves (DISTRONIC PLUS with steer assist), full stop and restart (Stop&Go Pilot), driver intervention, and a vehicle malfunction.

IV. QUALITATIVE RESULTS

Participant commentary from the interviews (pre-intro, post-intro, post-drive) and during the test drives was recorded. This Section includes excerpts from the commentary and participant behaviors recorded. The commentary was transcribed and categorized according to the TiA factors examined by the questionnaire [12], and additionally participant *Driving Style* and *Weather*.

A. Reliability/Competence

Toward the end of the drive, P9 said, "I think the benefits [of ADAS] are clear and undeniable. Every system here is intended to improve safety. I don't think there's any danger posed by the system," and "I was skeptical. Having seen it in action, having felt it under my hands…it is a good thing and I could recommend this kind of system to other people as well. I could talk positively about my experiences on the road. I wouldn't be averse to having this kind of system in my own vehicle."

P4 took back control while in an autobahn construction zone due to discomfort with Steer Assist, stating, "It's keeping us in the lanes but before it was a little bit problematic. It went too far to the right and then it went to the left and then I intervened because I was not sure if it would do it itself."

After the drive, P4 said, "There were a couple of mistakes [with Steer Assist] and it was not too clear to me if it was on or off. I guess that's not the point of the system, that I have to focus more on the [system] than the road. It's not as useful because I have to keep my hands at the wheel anyway." During the course of the test drives the system experienced one malfunction, which occurred during P4's test drive. While driving on a country road, the system drifted the vehicle out of its lane. P4 allowed the vehicle to continue drifting out of lane until half of the vehicle was in the lane of the oncoming traffic before intervening.

B. Understanding/Predictability

After the introduction to the ADAS, P7 stated, "I was curious and skeptical at the beginning but now that I know more about how [the systems] work and what they can do for me, it makes me more confident. I think I may struggle with using them due to a lack of experience. I can trust [assistance] more than a full take-over of my driving."

P7 expressed how unpredictable vehicle behaviors affected feelings of trust, "I felt that I was mostly in control of the system, but not when resuming my settings. I knew I was faster when I last used the feature, so I wanted to use it to accelerate. But sometimes it was much faster than I expected. It was alarmingly fast. I did not trust in the braking after such a strong acceleration."

1) Mental Model: Several participants made comments that revealed changes in their mental model of the system as the drive progressed.

During the test drive, when P5 activated the system, they did not release their foot from the pedal until ~40 minutes in to the drive, which automatically caused the distance control system (Distance Pilot) to become passive, leading P5 to become confused about the system's functionality.

P2 said, "I see the lane is not there, so I will not trust [the system.]" and "It is steering but I want to make sure I keep my hands on the wheel because there is no lane marking right there." P6 expressed, "Maybe the [steer assist] is off because it has to "take some information in to analyze the situation," indicating their perception of the system's functionality. P9 stated, "I thought maybe if I moved to the right a little bit, [the system] would start to see a pattern in the lanes. I was trying to show the car what the lane looks like. I thought maybe if you just adjust the position of the car within these lanes that it could find a pattern within it, that it could orient itself."

P8 mentioned a time they adjusted their mental model and hence behavior, "Most times I felt in control. Except for the two times when the car ahead [moved into the adjacent lane and] turned off the road. I thought the system misinterpreted it a bit. I anticipated it the second time, based on the first time, that it could happen, and my anticipation was right."

C. *Intention of the Developers*

Participants (P1, P3, P5, P7) said that they felt "safe and comfortable" in the vehicle due to the brand. Several mentioned that they would prefer to drive a vehicle from another brand (P1, P8, P9). After an introduction to the system, P9 said, "I wouldn't say I necessarily feel better about [the system]…I think the developers did their best to create a system that doesn't put people in harm. I have faith in them but there are so many variables on the road that I don't feel comfortable putting my full trust in the system."

a) Implicit Information: P9 referenced stories they have heard in the media: "On the news you'll see, in the United States in particular, where people are very excited about automated driving systems, that someone runs into something because they are not paying attention and then ends up in a fatal accident. That puts me on edge about the whole thing. I am not exploding with excitement [to use the system]."

D. *Propensity to Trust*

In the initial interview, P1 said, "I am like a dinosaur. I don't have a big trust in the system. I feel safe with things I can see." After an introduction to the system, P1 expressed their feelings towards driving: "If I drive, I am concentrated on it and I like it. I would not like a car to do things for me that I could do myself." P7 stated, "I am curious and also suspicious if it really works. I think that I can do better than the automatic calculations of the system. It is making me curious but also cautious."

E. *Trust in Automation*

P5 expressed a desire to test this system but at each opportunity they intervened and took back control; in the end, stating, "I can't trust the system so fast. You would have to put a wall out of cushions in front of me before I try that" (referring to the Stop&Go Pilot). After the test drive, P5 stated "It's up to me how much I trust the system. If I would drive with such a system for a long time, I would put more trust in it. But now it's a very big contrast in driving for me."

P6 said, "I am not totally trusting but for me a [verbal trust score of] 4 is very high because I normally do not like these systems. But it is very comfortable to me now."

Throughout the test drive P10 was animated and expressive stating at the end, "It is a bit creepy for me…to trust a car. Normally, you trust a driver. I hate to go by airplane. Because you have to trust someone else. A stranger! But here...you have to trust a car...something with no inside, no feelings! It is only a system, a machine, and you have to trust it. The longer you drive it and the more you get familiar with it…you get a feeling for it and you start to trust the system."

F. *Driving Style*

During the test drive, P9 reflected on the effect the ADAS might have on driving styles, "I could imagine the system really reducing reckless driving. I don't feel the need to even worry about passing this person. I kind of just feel comfortable letting the car takeover. It kind of takes the pressure off me to take some sort of an action. If driving manual transmission, I would probably be more aggressive right now."

P9 stated they would "feel better" if other cars around them were using ADAS. "There are a lot of really bad drivers. I would know the car is going to adjust to keep them within tolerance limits automatically. I would probably feel better about being in traffic with the person." P10 said, "Maybe I would pay more attention to a car [with ADAS]. Because I know it is new a technology. But you can expect more of what a system would do than a human. A system would work or not work, not be in-between like humans. Maybe these are the cars on the road now that you think, 'Oh that driver drives very correctly.'"

1) Risk-taking: Shortly after expressing their apprehension at the start ("I am very nervous about what we're going to do today"), P3 engaged in repeated attempts to test the system at high speeds while on the autobahn. P3 intentionally drifted over lane markings several times to see if the Active Lane Keeping system would reorient the vehicle properly in the lane.

20 minutes into drive, P8 "provoked" the system, stating, "Now [the system] steered, because I provoked it. I tried to go straighter than I should have into the turn. If you get the angle of the curve wrong, it's nice that someone assists you with it."

When testing the Stop&Go Pilot, P7 said, "The car will stop? Cautiously, I am trying that. I've got my foot over the pedal. The car…the car completely stopped. It is new and weird. Okay, wow. Now it's going again on its own. If I know this is doing the job for me, I feel comfortable in releasing my foot and not keeping it directly above the pedals." P9 stated, "So it's going to stop completely? That's giving me a little bit of apprehension right now. That feeling, do I let it? Cause that's like twenty years of driving experience inching up towards that bumper."

Several of the participants drove through or attempted to drive through roundabouts in urban settings while the system was active. P4 followed a vehicle in front through the roundabout and out of the second exit. P8 also followed a vehicle into the roundabout but intervened as the system accelerated once the vehicle in front exited. P6, P7, and P10 approached the roundabout with ADAS active but intervened.

P4 was the only participant who activated the ADAS near its top speed at 190 km/h.

a) Hands Off: All participants with the exception of P5 removed their hands from the wheel long enough to trigger the hands-off warning graphic and/or auditory warning tone.

Halfway through the drive, while traveling on the autobahn (160 km/h), P1 crossed their arms. P1 also adjusted the position of their headrest at high speed, stating "See, this is something I would do now, because the system is on" as they put both arms behind their head and adjusted the headrest.

P4 stated, "You get a warning to put your hands on, but you don't have to do anything, it's just going on its own anyway." P4 discovered a work-around for disabling the hands-off warning; by briefly nudging the steering wheel slightly from side to side they were able to cease the warnings temporarily. P4 continued to workaround these warnings, through a narrow construction zone.

P3 told the hands-off auditory warning to "shut up" and P4 referred to it as "annoying." P4 received the most hands-off warning notifications of all participants (over 45 notifications).

2) Risk-aversion: Participants P1, P2, P5 and P6 did not allow the Stop&Go Pilot to come to a full stop while all other participants did. P2 had the lowest top speed at 130 km/h and possessed the most cautious driving style.

G. Weather

One instance of rain occurred which lasted approximately 15 minutes toward the end of P3's test drive. P3 said, "I like the system. I trust the system. But because of the rain I have not such a safe feeling because I don't know this car and I am driving it for the first time. It's not like you just sit here and feel safe, because it's up to all of the things that can happen around you." At the end of the drive, P3 stated, "If there was no rain, I would give the system a [verbal trust score of] 5 because it worked, and it did what it was supposed to do. Because of the rain, I didn't feel so safe, so I will say 4."

V. Quantitative Results

A. Verbal Trust Scores

Participants were asked to give a verbal rating of trust in the vehicle's ADAS on a scale from 1 to 5 (1=low trust, 5=high trust) three times throughout the experiment session: pre-interview, post-introduction, and post-drive.

Verbal trust scores indicated that six of the participants had an increase in trust after receiving an introduction to system (ranging from +0.25 to +1.50 compared to pre-interview scores). Three showed no change and one reported a decrease in trust (-0.50). Post-Drive, seven of the participants reported an increase in trust (ranging from +0.50 to +2.50), while one reported no change and two reported a decrease in trust (-1.0). A non-parametric Friedman test of differences among repeated measures, adjusted for ties was conducted and rendered $\chi2(18)=0.07$, $p>0.05$, which was nearly significant (see Figure 1).

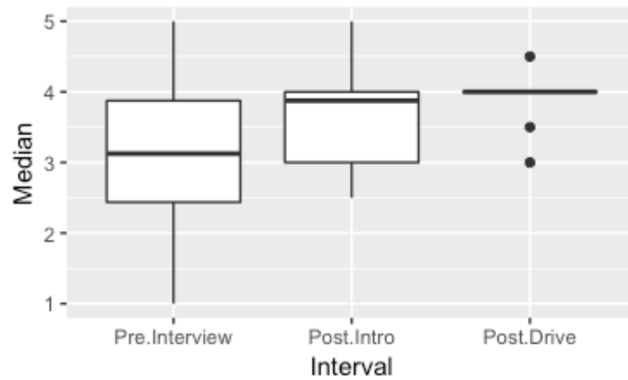


Figure 1. Comparison of median verbal trust scores for all participants at key intervals during the experiment, which were nearly significant at p>0.05.

B. TiA Questionnaire

The Trust in Automation questionnaire was administered twice during the experiment session: after participants were introduced to the system (pre-drive) and again after the test drive (post-drive). Scored results (see Section III.C.1. *Data Analysis*) from the questionnaire indicated that three participants had a decrease in TiA after the test drive (P1, P5, P4) while all other participants reported an increase of TiA after the test drive (see Table IV). All of the participants who reported a decrease in trust after the drive were >30 years of age with ten or more years of driving experience.

TABLE IV. TRUST IN AUTOMATION SCORES: PRE-DRIVE VS POST-DRIVE

Participant	Interval		
	Pre-Drive	***Post-Drive***	***Change in TiA***
P9	57.408	82.432	+ 25.024
P8	54.464	72.128	+ 17.664
P10	45.632	61.824	+ 16.192
P7	47.104	63.296	+ 16.192
P2	64.8	75.1	+ 10.3
P3	63.3	72.1	+ 8.8
P6	57.408	61.824	+ 4.416
P1	50	47.8	- 2.2
P5	75.1	63.3	- 11.8
P4	54.5	35.3	- 19.2

By participant, pre-drive and post-drive. An increase in TiA occurred in all participants while a decrease in TiA was observed in P1, P4 and P5.

In one instance, a participant's (P1) verbal trust scores did not align with their self-reported TiA responses. P1 verbally reported a gain in trust post-drive, but in the post-drive questionnaire reported a loss of trust.

Each participant's TiA response (based on the 5-point Likert scale) was recoded, and a pre-drive median and post-drive median value was given for each participant. A Wilcoxon Signed-Rank Test performed on all participant's pre-drive and post-drive medians indicated that the post-drive TiA median scores were not significantly higher than pre-drive median TiA scores (*Z*=-0.86, p>0.05) (Figure 2, factor: all).

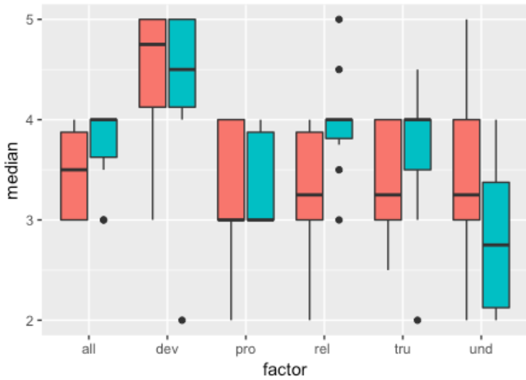


Figure 2. Trust in Automation, pre-drive (pre) & post-drive (post) medians. Shown overall (all) and for each factor (dev=*Intention of the Developers*, pro=*Propensity to Trust*, rel=*Reliability/Competence*, tru=*Trust in Automat*ion, und=*Understanding/Predictability*).

The factors *Reliability/Competence*, *Understanding/ Predictability*, *Intention of the Developers*, *Propensity to Trust*, *Trust in Automation* were also considered individually for analysis. The post-drive median score for *Reliability/Competence* was found to be significantly higher than the pre-drive median score (*Z*=-2.17, p<0.05) (see Figure 2, factor: rel). The pre-drive vs. post-drive median scores for the other factors were not found to be statistically significant at p>0.05.

C. Facial Emotion Recognition

The driver facing camera footage for each participant's test drive was processed by the convolutional neural network for FER. A value ranging from 0 to 1.0 for each emotional state (where each of the four emotions, *happy*, *angry*, *surprise* and *neutral* share a portion of a 1.0 value) were returned every one tenth of a second for the entirety of the drive. FER scores were calculated for the entire duration of each participant's test drive, which lasted approximately one hour. The visualization of each journey over time was plotted in MATLAB. Frames in which no face was detected were included in the analysis with *movmean* and the *dim* property set to 5,000 (smoothing) (Figure 5, see next page). Each graph is labeled with the respective participant number, their FER score for four emotions out of 1.0 is shown in the y-axis. As shown in the legend below the plots, *neutral* is plotted in blue, *happy* is plotted in red, *surprise* is plotted in yellow, and *angry* is plotted in purple. The x-axis plots location in time, in which *Start* marks the beginning of the test drive and *End* marks the end of the test drive. Entry (EA) and exit of the autobahn (XA) are also annotated (see the caption of Figure 5).

D. Facial Emotion Recognition + Sample Events

A selection of driving events was timestamped and extracted from FER analysis and the visualization of each event was plotted with movmean in MATLAB with the *dim* property set to 100 (smoothing). The selection here represents only a portion of the driving events which occurred. As shown in the legend for each graph, *neutral* is plotted in blue, *happy* is plotted in red, *surprise* is plotted in yellow, and *angry* is plotted in purple.

1) Hands Off

While the Mercedes-Benz ADAS used in this study is not a hands-free system, a majority of participants elected independently to remove their hands from the wheel at some point during the test drive. The steering wheel was equipped with sensors (Mercedes-Benz, standard production) as a means of measuring driver attentiveness. When the system detects a lack of contact with the steering wheel, the system will display a graphic warning indicating the driver should return their hands to the wheel. If this warning is not heeded, it is then accompanied by an auditory tone. Figure 3 highlights the emotional state of the participant as they experiment with removing their hands from the wheel for the first time.

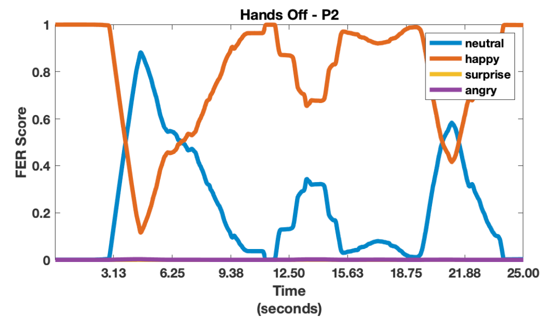


Figure 3. Hands Off, Participant 2. At 3.33 seconds the driver removed their hands from the wheel. At 10-15 seconds, they stated, "This is interesting. It keeps the [lane] all by itself." At 21 seconds, the driver said, "Hmm. Crazy."

1) Deceleration

When active in an autobahn setting, DISTRONIC PLUS (Mercedes-Benz adaptive cruise control) is capable of autonomously decelerating the vehicle at high speeds. Figure 4 and Figure 6 are samples of participants emotional state as the vehicle decelerated autonomously.

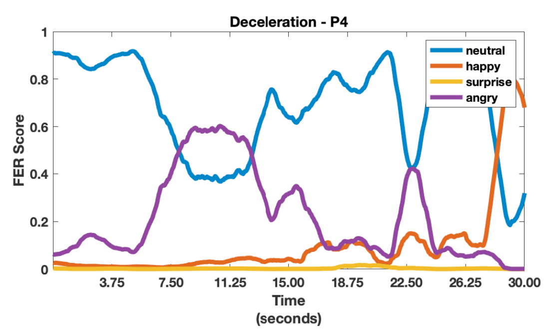


Figure 4. Deceleration, Participant 4. While on the autobahn with ADAS active near top speed (190km/h), the vehicle slowed itself autonomously, approximately 50% to 90km/h. The vehicle begins decelerating at 4 seconds, and the driver says, "Ok!" At 13 seconds the driver says, "That was not too comfortable for me, but I guess that's because my car doesn't brake so hard." At 23 seconds they state, "I would have braked way before that."

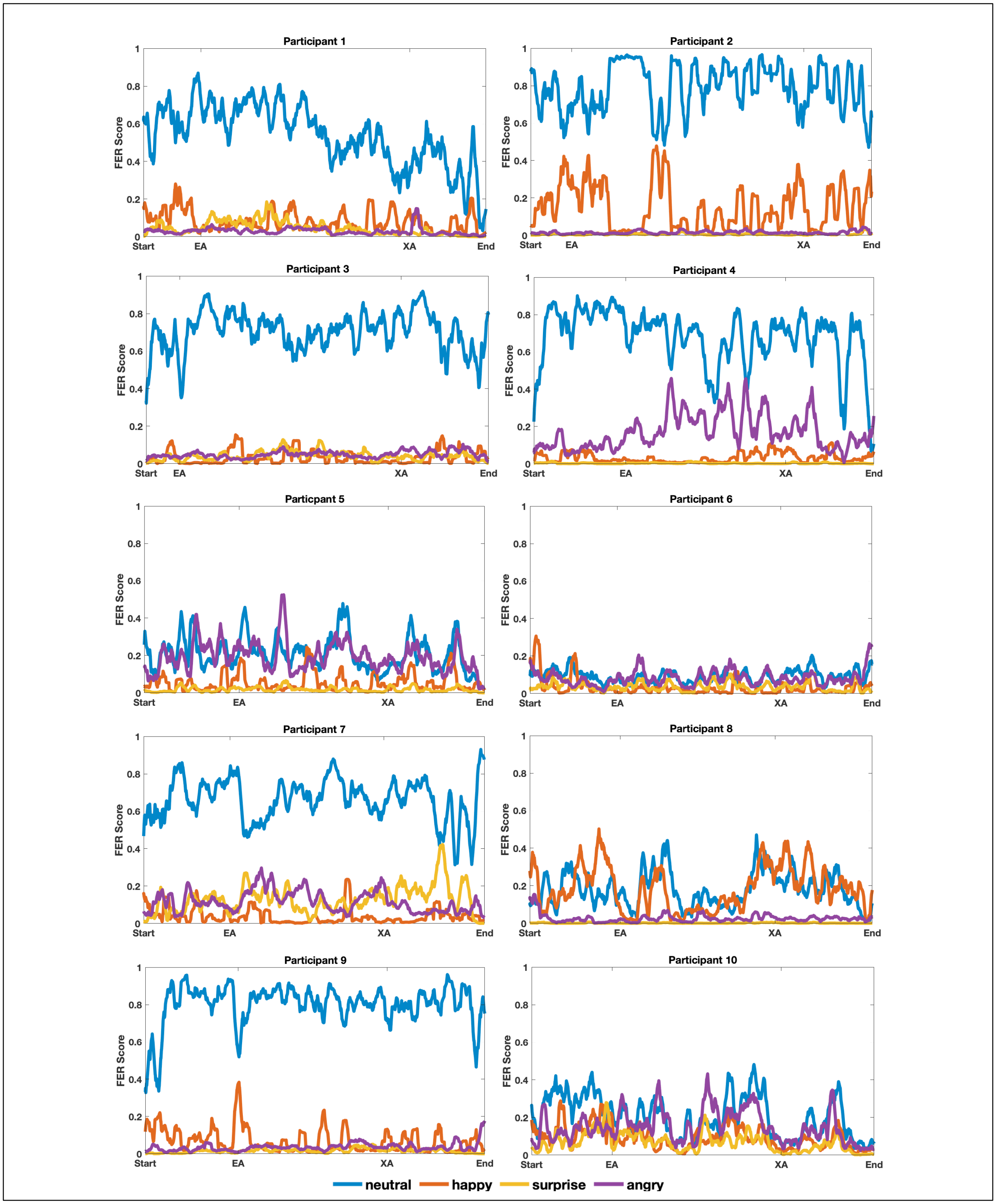


Figure 5. Facial Emotion Recognition Scores for all participants' entire test drive session. Y-axis expresses participants' FER scores from 0-1.0 where all emotions share a portion of 1.0. X-axis shows location in time, in which *Start* marks the beginning of the test drive session, *EA* marks "Enter Autobahn", *XA* marks "Exit Autobahn", and *End* marks the end of the test drive session. Between *Start* and *EA*, *XA* and *End*, participants drove in country road and urban settings. Between *EA* and *XA* participants drove on the autobahn. Test drive sessions averaged 70 minutes in total. P5, P6, P8, and P10 wore corrective eyeglasses which impeded facial detection. This is visible in the relatively low FER scores of these participants (see Section VI for more).

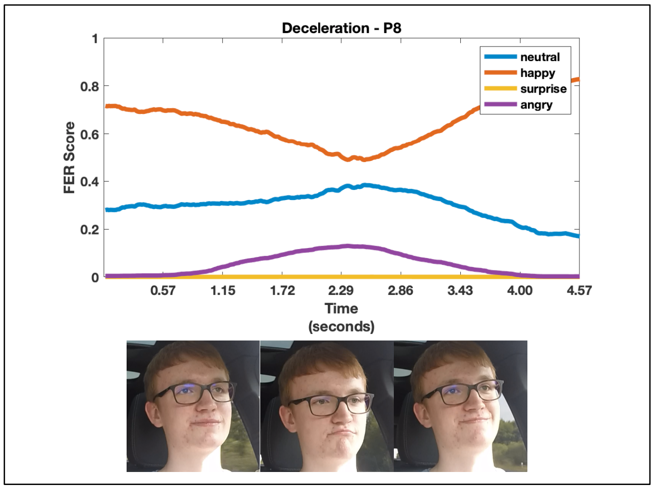


Figure 6. Deceleration with photos of corresponding frames, Participant 8. While traveling on the autobahn with ADAS active, the vehicle braked autonomously as it approached the vehicle in front. The photo on the left was just before 2.5 seconds, the middle at 2.5 seconds and the third after.

2) Steering

DISTRONIC PLUS with Steering Assist is capable of autonomously handling light curves with the supervisor of the driver. Figure 7 highlights the emotional state of the driver as they experience the vehicle steering through a curve autonomously for the first time.

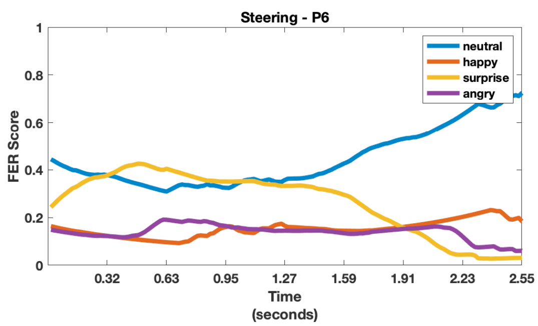


Figure 7. Steering, Participant 6. The curve of the FER score for neutral mimics the physical curve the driver was passing through on-road (as *neutral* and *happy* rises, the vehicle is coming out of the curve).

3) Intervention

Participants elected to turn off the ADAS and take back full control of the vehicle at different times or different reasons. This includes events in which the driver felt that the vehicle was not acting quickly enough, for example: DISTRONIC braking occurring to slowly, DISTRONIC acceleration occurring too quickly/slowly, uncertainty and/or distrust that the vehicle would come to a complete stop on its own (Stop & Go). Figure 8 shows the emotional state of Participant 7 as they elected to intervene and take back control.

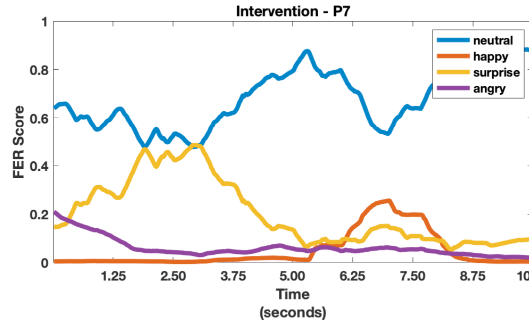


Figure 8. Intervention, Participant 7. In a country road setting, the vehicle began to slow slightly as it was approaching a vehicle in front turning off the road, however the driver intervened between 1.66-3.33 seconds. At 6.66 seconds the participant said, “I took over, it was me.” They explained that they felt vehicle was not stopping quickly enough.

4) Malfunction

During the test drive sessions, while the ADAS was active, one obvious malfunction occurred. While active on a country road, the vehicle drifted out of lane and halfway into the lane of on-coming traffic. Figure 9 shows the emotional state of Participant 4 as they silently corrected this error.

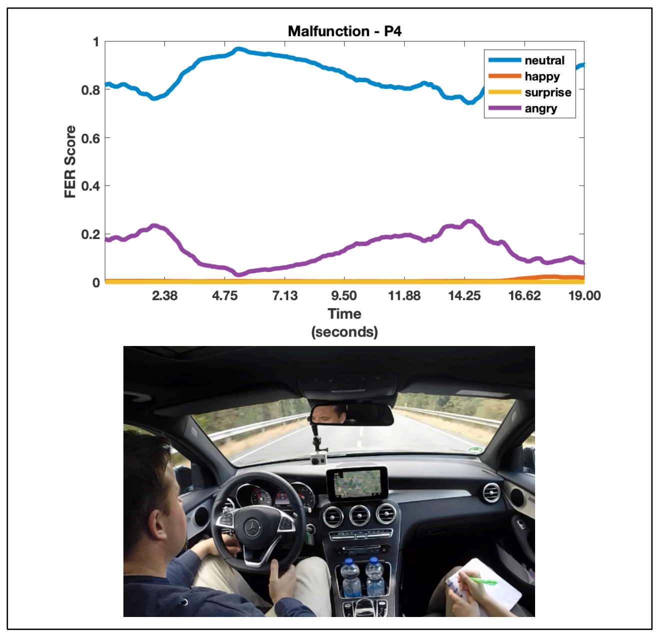


Figure 9. Malfunction with corresponding photo depicting vehicle having drifted into the lane of on-coming traffic (at 6 seconds), Participant 4. Interestingly, the driver remains neutral at the time of the event, with slight rise in anger following, continued by a very slight detection of *happy*.

5) Stop&Go

A portion of the participants elected to test the Stop & Go feature of the system, which allows the vehicle to come to a complete stop. This feature was used while in traffic in autobahn and urban settings. Figure 10 highlights the first time Participant 7 used this feature in an urban setting.

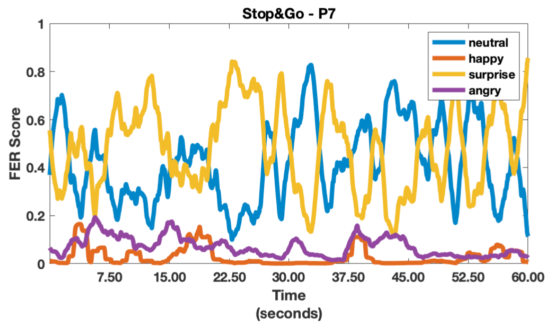


Figure 10. Stop & Go, Participant 7. While in an urban setting P7 allows the vehicle to come to a full stop. At 10 seconds, they say in anticipation, "The car is braking" and then, "I [will] try to trust the car and use the feature as it's offered. I am not doing anything." The vehicle stops at 30.83 seconds. At 35 seconds, because the car was stopped for less than three seconds (the threshold for the driver having to intervene from a full stop), the vehicle accelerated on its own to follow the car in front.

E. Facial Emotion Recognition + TiA Score

The overall mean values for each of the four emotions were noted separately for each participant. Values were then converted into a percentage, indicating which emotions were most dominant throughout each drive for each participant, respectively. Figure 11 displays the relationship between participant's reported TiA scores and FER scores. The y-axis reflects the change in participants pre-drive vs. post-drive TiA scores, in order from the greatest gain to the greatest loss in TiA. The x-axis presents the FER score as a percentage, indicating the dominant emotion for each participant's drive. *Neutral* was the dominant emotion among all participants, however participants displayed differing frequencies of the emotions *happy*, *angry* and *surprise*. Participants with a gain in TiA post-drive tended to display *happy* whereas participants with a loss in TiA post-drive tended to display *angry* (see Figure 11). Note the distribution of angry among participants and its prevalence in those with a loss in TiA.

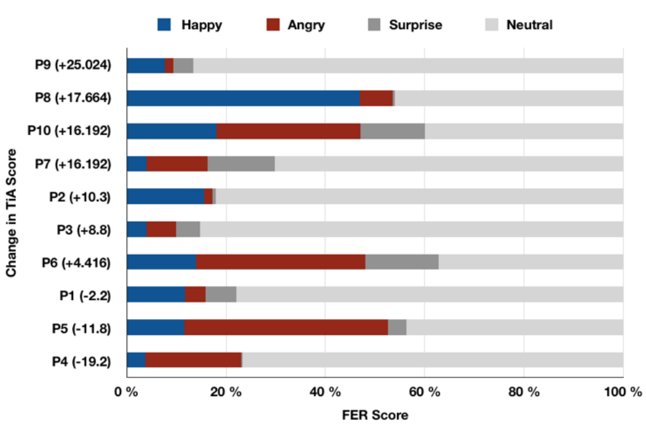


Figure 11. Participant change in TiA Score (y-axis) compared with FER scores (x-axis) as a percentage of their total test drive session.

According to their TiA scores, participants P9 and P8 reported the greatest gain in trust and (with the exception of *neutral*) displayed *happy* as a dominant emotion. Participants P4 and P5 reported the greatest loss of trust and (with the exception of *neutral*) displayed *angry* as a dominant emotion. A loss of trust was observed in both P4 and P5's verbal trust scores, which decreased at the same interval (both -1.0 post-drive), while P8 and P9 reported an increase in trust at the same intervals (both +1.25 post-introduction and post-drive).

P7 and P10 reported the same gain in TiA post-drive yet displayed different dominant emotions according to FER. A comparison of the verbal trust scores of P7 and P10 do not reveal the same changes in trust (P7 reported verbally no change in trust post-drive, while P10 reported +1.0 post-drive). Additionally, P10 displayed the dominant emotions as those drivers which had the greatest loss in trust (P4 & P5).

Relative to the other drivers in this study, P4 possessed the most aggressive driving style (e.g., speed, triggering of hands off warning graphic/tone). A contrast to P5, who was the most reluctant to give over control to the system and did not take their hands-off the wheel. However, both P4 and P5's TiA scores revealed the greatest loss of in trust in automation post-drive. This loss of trust is also reflected in their verbal trust scores, as they were the only participants to verbally report a decrease in trust post-drive. Further, P4 and P5's both displayed *angry* as a dominant emotion.

VI. Discussion

Nesting automation in safety-critical systems requires careful consideration from a human-machine interaction perspective. In order to determine what effect a user's first contact with an ADAS has on their level of trust in the system, an on-road experiment with ten participants with no prior experience with ADAS was conducted. The use of the Trust in Automation questionnaire [20], verbal trust scores, Facial Emotion Recognition, and interviews/observational data, enabled a mixed method analysis of each participant's experience. It was hypothesized that participants would report higher levels of trust in ADAS after their first experiential drive, and that FER results would reveal a relationship between a participant's TiA score and the emotions displayed (*happy*, *angry*, *surprised*) during the drive. Further, an exploratory analysis of participant's FER scores, over the course of their test drive session and during a selection of driving events was conducted.

The scored results of the TiA questionnaire revealed that trust in automation increased after the test drive in a majority of the participants. A comparison of pre- vs. post-drive median TiA scores however, did not reveal a statistically significant difference in trust in automation ($p>0.05$).

The significant rise in factor *Reliability/Competence* ($p<0.05$, Figure 2) after the drive indicates that based on their experience driving with the ADAS, participants believe that the system performed in a way that reliably assisted them in achieving their goals [4]. Perhaps inherent trust in the established Mercedes-Benz brand had an effect on the initial scores, but it is not clear what effect it might have had on the scores post-drive. Additionally, participants made comments corresponding to the underlying factors of trust in automation, for example, P1 referred to themself as a "dinosaur" regarding their approach to technology, which may be interpreted as an indicator of their *Propensity to Trust*. Reviewing P1's median TiA score for the factor *Propensity to Trust* reveals a low score (pre-drive: 3, post-drive: 2). P9 mentioned their, "faith in the developers" during the experiment session. P9's median TiA score for the factor *Intention of the Developers* was high (pre-drive: 5, post-drive: 5).

As noted, a system malfunction occurred during the later part of P4's test drive. One can assume that this was weighed as a factor in their reported feelings of trust in the system. Additionally, it is plausible that the *Hawthorne Effect* [26] may have occurred in the instance where a participant (P1) reported an increase in trust verbally post-drive, but a decrease in trust on the questionnaire post-drive.

Regarding FER scores, participants with a gain in Trust in Automation post-drive tended to display *happy* more frequently in their FER score while those with a loss in TiA post-drive tended to display *angry* more frequently (Figure 11). This finding is of interest, as self-reported feelings of trust in automation and emotional states appear to follow a similar pattern. This gives validity to the combination of TiA, verbal trust score and FER data, suggesting that this approach may be able to identify a specific persona, who may be less trusting and therefore less accepting of ADAS. However, due to some discrepancies (see Section V.E), additional research is needed to determine if a relationship between trust in automation and emotions captured via FER can be replicated.

The driving behaviors of the participants demonstrated a willingness to take risks with the system, for example: using Stop&Go Pilot, hands-off events, and attempting to use the system in complex scenarios such as construction zones, roundabouts, and urban settings. This is aligned with the definition of trust by Mayer et al. [11]. Based on the results of this study however, displaying a willingness to take risks with the system alone is not a reliable indicator of trust as the participant who took the most risks with the system (P4) reported the greatest loss of trust post-drive.

Referring back to the LCoT framework by Ekman et al. [17], the *Learning Phase* events, *Control Transition 1, Automated Mode* and *Control Transition 2* (handover scenarios) do not list *Mental Model* as a trust-affecting factor. This is contradicted by the observations in this study. For example, during the test drive, while thinking aloud, participants stated their beliefs about how the system would behave prior to engaging in *Control Transition 1*. While in *Automated Mode* participants stated their expectations of the system's behavior. When the system behaved in a way that was not aligned with their expectations, participants engaged in *Control Transition 2*. Participants then stated why they took back control, and based on their learning from the scenario, adjusted their mental model to adapt their future interaction with the system (see Section IV.B.1). Participants who more easily developed an accurate mental model aligned with the functionality of the system handed control over to the system more easily whereas those whose mental model was not well aligned with the system had a difficult time handing over control to the vehicle. For example, P5 was reluctant to release their foot from the accelerator for an extended period of time, indicating they possessed a poor mental model of the system's functionality. After several *Transitions*, P5 made adjustments to their mental model and their interaction became more fluid. In contrast to P9, who gained an understanding of the system functionality quickly, expressed a desire to work with the system by showing "the car what the lane(s) looks like." P9's accurate model of the system (that the vehicle is tracking the lane markings) allowed for them to place more trust in the system. This suggests that during the *Learning Phase,* both *Automated Mode* and *Control Transition* events are impacted by the trust-affecting factor *Mental Model*.

This study confirms Ekman et al.'s [18] conclusion that a mixed methods approach is required to understand trust in automation. Results also suggest that the finding by Gold et al.'s [19] simulator study ("driving experience increased self-reported trust in automation") does in fact carry over to the on-road context of use.

An exploratory analysis of facial emotion recognition data captured via the driver facing camera from the on-road test drives was conducted to investigate which emotional states accompanied the participants total test drive session and assisted driving events (e.g., assisted steering). The selection of driving events and their corresponding FER plots illustrate that driver emotional states during specific driving events can be quantified by means of FER.

Analysis of FER for the whole of the participant's test drive revealed that participants who wore corrective eyeglasses (P5, P6, P8, P10) had lower FER scores. Further analysis of the raw FER data indicated that this was due to the FER algorithm having increased difficulty in accurately detecting the faces of participants who wore corrective eyeglasses. Frames with an undetected face ranged from 20% missing (P5) up to 78% missing (P6). A post-hoc review of literature on eyeglass detection for FER revealed that this issue is not uncommon [39] and CNN techniques for improving detection [40], [41]. Perhaps further training of the CNN on a broader dataset which includes more domain-specific footage (i.e. images of drivers, and drivers wearing eyeglasses) may be the best approach to improving detection for this domain.

Still, FER plots in Figure 5 are challenging to compare between participants. A variety of factors unrelated to the act of driving may have affected a participant's FER score. For example, driving events occurred in differing durations and frequencies, and participants may have been discussing a prior driving event while engaging in another driving event (e.g., discussing a time they had to intervene while their hands were off of the wheel, as participants were encouraged to think aloud). Other factors impacting FER score may have been things like sunny weather, which may have caused a driver to squint, perhaps causing them to appear *angry*, when in fact they were not. A participant's individual culture or personality may also play a role in the way they display emotions. Participants' accompaniment by a moderator in the passenger seat may have also affected the emotions displayed.

This raises an important consideration regarding affective computing in-vehicle. Observations of participants' emotional state fluctuating in response to vehicle behaviors may in the future enable the calibration of system behaviors to meet user expectations of the system's response to the on-road environment in a development setting. However, often while the ADAS was active during extended periods of autobahn driving, participants would begin a conversation on topics unrelated to the driving environment, becoming expressive while talking about past or future events. One can assume that in conditional or highly automated driving (in which drivers are able to be fully out of the loop) that users will display emotions correlated with non-driving related tasks versus the actual on-road environment or system interactions. While FER may be a useful tool for researchers & system developers, FER may not be useful for affective computing in-vehicle at the consumer level.

VII. LIMITATIONS

Studies in simulators and on closed courses allow for significant control over research conditions, whereas studies on public roads leave much open to chance. Due to the spontaneity of the on-road environment, each participant was exposed to a variety of different scenarios at varying frequencies and intervals throughout the drive, with unplanned events such as rain or a system malfunction occurring simply by chance. Furthermore, because there is no baseline or reference point for each participant's emotional state in this study (e.g., FER scores for a drive along the same route in their daily driving vehicle), one should be cautious in their interpretation of this data. Participants also may have had a different FER scores had they been driving alone and not accompanied by a moderator.

While this study provides insight into the development of trust in ADAS on-road and using FER "in the wild", one should be cautious in generalizing the results of this study. The sample size was small (n=10), as it was limited by the loan period of the vehicle supplier. Further, the results of this study are specific to the design of Mercedes-Benz ADAS, and experienced or daily drivers of this system may display different emotions under similar conditions.

VIII. CONCLUSION AND FUTURE WORK

An enhanced understanding of the exchange between the user and the system on-road and the resulting effects on trust, will aid in the design of safer, more efficient automated systems. A method for correlating FER data and TiA scores is presented here which may be explored in future studies. FER accuracy in terms of both detection and classification could be improved by means of data augmentation or by training the network on a more robust dataset, should it be used again in the future. Running further test drives without a moderator in vehicle may result in different FER scores. Analysis with an emotion recognition method specifically for driver speech may also be insightful, especially with the movement towards voice user interfaces in-vehicle. The finding that all participants who reported a decrease in trust after the drive came from a similar demographic (>30 years of age, ten or more years of driving experience) warrants further investigation.

The results presented here support future research of trust in vehicle automation and other applications of FER in-vehicle. More research is needed to improve the understanding of the development of trust in automation, in order to aid the user in their acceptance of this safety-critical technology. Tackling issues of trust in ADAS today lays the groundwork for the acceptance of higher levels of autonomy in the future, eventually leading to fewer deaths, less injury, disability and a safer more enjoyable on-road experience for all.

ACKNOWLEDGMENTS

Thank you to Herbrand Mercedes-Benz and Mr. Sven Ingenpaß for the generous loan of the vehicle for this project. To the administration and staff of Hochschule Rhein-Waal University of Applied Sciences for their support and to the university staff who participated in this study, thank you. Special thanks to Dr. Claudio Abels for his assistance with logistics, Dr. André Frank Krause for his assistance on the CNN for FER, and Ms. Sabine Lauderbach for her knowledge of statistics.

REFERENCES

[1] L. Dixon, W. M. Megill, and K. Nebe, "Trust in Automation : An On Road Study of Trust in Advanced Driver Assistance Systems," in *VEHICULAR 2019 : The Eighth International Conference on Advances in Vehicular Systems, Technologies and Applications*, 2019, no. c, pp. 85–93.

[2] World Health Organization, "Global status report on road safety 2018," World Health Organization, 2019.

[3] C. Kerry and J. Karsten, "Gauging investment in self-driving cars," *The Brookings Institution*, 2017. [Online]. Available: https://www.brookings.edu/research/gauging-investment-in-self-driving-cars/. [Accessed: 09-Oct-2019].

[4] J. Lee and K. See, "Trust in Automation: Designing for Appropriate Reliance," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 46, no. 1, pp. 50–80, 2004.

[5] Society of Automotive Engineers (SAE) International, "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles: J3016_201806," 2018.

[6] H. Abraham, B. Seppelt, B. Mehler, and B. Reimer, "What's in a Name : Vehicle Technology Branding & Consumer Expectations for Automation," *Proc. 9th ACM Int. Conf. Automot. User Interfaces Interact. Veh. Appl. (AutomotiveUI '17)*, pp. 226–234, 2017.

[7] S. Trösterer *et al.*, "What We Can Learn from Pilots for Handovers and (De)Skilling in Semi-Autonomous Driving: An Interview Study," *Proc. 9th Int. Conf. Automot. User Interfaces Interact. Veh. Appl. (AutomotiveUI '17)*, pp. 173–182, 2017.

[8] N. Hutchins and L. Hook, "Technology Acceptance Model for Safety Critical Autonomous Transportation Systems," *2017 IEEE/AIAA 36th Digit. Avion. Syst. Conf.*, pp. 1–5, 2017.

[9] C. Rödel, S. Stadler, A. Meschtscherjakov, and M. Tscheligi, "Towards Autonomous Cars: The Effect of Autonomy Levels on Acceptance and User Experience," *Proc. 6th Int. Conf. Automot. User Interfaces Interact. Veh. Appl.*, pp. 1–8, 2014.

[10] M. Nees, "Acceptance of Self-driving Cars: An Examination of Idealized versus Realistic Portrayals with a Self- driving Car Acceptance Scale," *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 60, no. 1, pp. 1449–1453, 2016.

[11] R. Mayer, J. Davis, and D. Schoorman, "An Integrative Model of Organizational Trust," 1995.

[12] M. Körber, "Theoretical considerations and development of a questionnaire to measure trust in automation," in *20th Triennial Congress of the IEA*, 2018.

[13] A. Miramontes *et al.*, "Training Student Air Traffic Controllers to Trust Automation," *Procedia Manuf.*, vol. 3, pp. 3005–3010, 2015.

[14] F. Eyben *et al.*, "Emotion on the Road—Necessity, Acceptance, and Feasibility of Affective Computing in the Car," *Adv. Human-Computer Interact.*, pp. 1–17, Jul. 2010.

[15] J. Izquierdo-Reyes, R. A. Ramirez-Mendoza, M. R. Bustamante-Bello, J. L. Pons-Rovira, · Jose, and E. Gonzalez-Vargas, "Emotion

recognition for semi-autonomous vehicles framework," *Int. J. Interact. Des. Manuf.*, 2008.

[16] Affectiva Inc., "Affectiva Automotive AI: Metrics in Affectiva Automotive AI," 2018. [Online]. Available: https://www.affectiva.com/product/affectiva-automotive-ai/. [Accessed: 17-Sep-2018].

[17] F. Ekman, M. Johansson, and J. Sochor, "Creating Appropriate Trust in Automated Vehicle Systems: A Framwork for HMI Design," *IEEE Trans. Human-Machine Syst.*, 2017.

[18] F. Ekman and M. Johansson, "Understanding Trust in an AV-context : A Mixed Method Approach," *Proc. 6th Humanist Conf.*, no. June, pp. 13–14, 2018.

[19] C. Gold, M. Körber, C. Hohenberger, D. Lechner, and K. Bengler, "Trust in Automation – Before and After the Experience of Take-over Scenarios in a Highly Automated Vehicle," *Procedia Manuf.*, vol. 3, no. November, pp. 3025–3032, 2015.

[20] M. Körber, E. Baseler, and K. Bengler, "Introduction matters: Manipulating trust in automation and reliance in automated driving," *Appl. Ergon.*, vol. 66, no. January, pp. 18–31, 2018.

[21] J. Jian, A. Bisantz, and C. Drury, "Foundations for an Empirically Determined Scale of Trust in Automated Systems," *Int. J. Cogn. Found. an Empirically Determ. Scale Trust Autom. Syst.*, no. January 2015, pp. 37–41, 2000.

[22] B. Ko, "A Brief Review of Facial Emotion Recognition Based on Visual Information," *Sensors*, vol. 18, no. 2, p. 401, 2018.

[23] T.-K. Tews, M. Oehl, F. W. Siebert, R. Höger, and H. Faasch, "Emotional Human-Machine Interaction: Cues from Facial Expressions," Springer, Berlin, Heidelberg, 2011, pp. 641–650.

[24] R. W. Picard, "Affective Computing," Cambridge, MA, 1995.

[25] H. Hastie, X. Liu, and P. Patron, "Trust Triggers for Multimodal Command and Control Interfaces," in *Proceedings of the 19th ACM International Conference on Multimodal Interaction (ICMI'17)*, 2017, pp. 261–268.

[26] L. Bortolotti and M. Mameli, "Decpetion in Psychology: Moral Costs and Benefits of Unsought Self-Knowledge," *Account. Res.*, vol. 13, no. 3, pp. 1–20, 2006.

[27] Daimler AG, "Active safety: Intelligent Drive: Assistance in all driving situations," *Daimler Global Media Site*, 2018. [Online]. Available: https://media.daimler.com/marsMediaSite/en/instance/ko/Active-safety-Intelligent-Drive-Assistance-in-all-driving-situations.xhtml?oid=10001778. [Accessed: 14-Sep-2018].

[28] J. Brooke, "SUS - A quick and dirty usability scale," *Usability Eval. Ind.*, vol. 189, no. 194, pp. 4–7, 1996.

[29] Lund Research Ltd, "Wilcoxon Signed Rank Test in SPSS Statistics," *Laerd Statistics*, 2018. [Online]. Available: https://statistics.laerd.com/spss-tutorials/wilcoxon-signed-rank-test-using-spss-statistics.php. [Accessed: 12-Oct-2018].

[30] M. Berenson, D. Levine, and T. Krehbiel, *Basic Business Statistics: Concepts and Applications*, New Jersey. Upper Saddle River: Prentice Hall, 2012.

[31] Statistics How To, "Friedman's Test / Two Way Analysis of Variance by Ranks," *Statistics How To*, 2014. [Online]. Available: http://www.statisticshowto.com/friedmans-test/. [Accessed: 20-Sep-2018].

[32] RStudio, "RStudio: Integrated development environment for R." Boston, MA, 2018.

[33] R Core Team, "R: A language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria, 2018.

[34] Felipe de Mendiburu, "agricolae: Statistical Procedures for Agricultural Research." R package version 1.2-8 , 2017.

[35] H. Wickham, "ggplot2: Elegant Graphics for Data Analysis." Springer-Verlag, New York, 2016.

[36] I. J. Goodfellow *et al.*, "Challenges in Representation Learning: A report on three machine learning contests."

[37] MathWorks Inc., "MATLAB." 2018.

[38] The Mathworks Inc., "Moving mean - MATLAB movmean R2018b," 2018. [Online]. Available: https://www.mathworks.com/help/matlab/ref/movmean.html. [Accessed: 06-Nov-2018].

[39] A. Bernin *et al.*, "Towards more robust automatic facial expression recognition in smart environments," *ACM Int. Conf. Proceeding Ser.*, vol. Part F128530, pp. 37–44, 2017.

[40] A. M. Basbrain, I. Al-Taie, N. Azeez, J. Q. Gan, and A. Clark, "Shallow convolutional neural network for eyeglasses detection in facial images," in *2017 9th Computer Science and Electronic Engineering Conference, CEEC 2017 - Proceedings*, 2017, pp. 157–161.

[41] A. M. Basbrain, J. Q. Gan, A. Sugimoto, and A. Clark, "A Neural Network Approach to Score Fusion for Emotion Recognition," in *2018 10th Computer Science and Electronic Engineering Conference, CEEC 2018 - Proceedings*, 2018, pp. 180–185.

International Journal on Advances in Intelligent Systems, vol 13 no 1 & 2, year 2020, http://www.iariajournals.org/intelligent_systems/

Knowledge Resource Object Development and Mining with a Knowledge-centric Architecture

Claus-Peter Rückemann
Westfälische Wilhelms-Universität Münster (WWU), Germany;
Knowledge in Motion, DIMF, Germany;
Leibniz Universität Hannover, Germany
Email: ruckema@uni-muenster.de

***Abstract*—This paper presents a knowledge resources' view on object development based on the new architecture framework, which is a research result of a series of practical problem solving implementations and further developments based on common knowledge concepts and integrated application components. The framework is considered knowledge-centric, supported by knowledge resources, which constitute the fundamental base and imply the core of key assets. Besides further knowledge development, the knowledge-centric architecture flexibly allows implementations of mining and computation components for many scenarios and the employment of available computation infrastructures. An important quality of the architecture framework is the inherent property to assign different roles for the professional tasks in creation and development cycles. These roles address the major complements of knowledge, including factual, conceptual, and procedural components as well as documentation. This paper discusses object development and knowledge mining based on a knowledge-centric architecture. The main goal of this research is to illustrate the universal multi-disciplinary and multi-lingual knowledge and content related features for major use-case activities, deploying a knowledge-centric architecture and to provide new examples from the practice of systematical object development and the implementation of flexible methods.**

Keywords–Knowledge Resource Object Creation; Object Development and Integration; Containers and Collections; Knowledge-centric Architecture; Knowledge-centric Mining.

I. Introduction

From information science view, knowledge-centric complements are in focus when creating and employing knowledge resources for mining purposes. Creation and further development of these resources and objects is of primary interest. This extended research is based on the knowledge-centric computation architecture and its deployment for knowledge mining, which was presented at the INFOCOMP 2019 conference in Nice, France [1]. The presented research concentrated on the new architecture and facilities to implement and describe activity groups in a consistent and standardised way.

Consequently, this extended paper concentrates on research from the knowledge resources' perspective, the object development, based on major use-cases groups, namely knowledge creation and development on the one hand and knowledge mining on the other hand. The development of knowledge resources and objects essentially contributes to the long-term values created for universal knowledge. Fostering structured organisation of content, which can be addressed by manifold methods and which can carry universal references to multi-disciplinary and multi-lingual knowledge is considered a primary benefit for development, valorisation, and flexibility of re-use.

The continuous development of knowledge resources and objects should be provided a number of major features, which are supported by the architecture, especially:

- Development of resources and objects,
- flexible object structures,
- object groups, e.g., collections and containers,
- multi-disciplinary knowledge object support,
- flexible object and context references,
- support of knowledge complements,
- multi-lingual features,
- consistency of resources,
- object re-use,
- knowledge mining,
- creation and deployment of algorithms,
- object integration,
- a fundamental, universal framework with a
- high potential of long-term facilities.

The goal of this research is to implement resources and features according to the knowledge-centric architecture. Use-cases illustrate and summarise the knowledge and content related features for major activities and to summarise the architecture and the implementation fundaments. The paper introduces major use-cases of object development and knowledge mining based on the architecture and presents relevant results for objects and content. The resulting object development based on this research is presented and demonstrated for a consistent view on objects, closely referring to activities for these use-cases.

This paper is organised as follows. Section II and III deliver motivation, fundaments and previous work. Section III presents the architecture resulting from this research. Sections IV and V introduce to the knowledge-centric architecture and to an implementation of two major use-cases for object development and mining. Section VI presents a resulting object development architecture. Sections VII and VIII present the respective results of object development and knowledge mining, illustrated by practical examples. Section IX discusses the main results and evaluates them in context of the practical application scenario. Section X summarises the results and lessons learned, conclusions, and future work.

II. Motivation

All implementations of mathematical machines, which we call 'computer systems' today, can strictly only deal with formal systems. Knowledge is a capability of a living organism and can itself not be incorporated by formal systems. Neither can intrinsic meaning, which is an essential characteristics of real knowledge and a unique stronghold of knowledge, be a matter of formal systems nor can mathematical relations, the theory of sets, exlusiveness or creating completeness be applied to knowledge.

Solutions requiring a wide range of knowledge content as well as implementations of algorithms and components are often challenging to handle, the more when it comes to operating and adding to the resulting solutions for decades or even further developing content and implementations for long-term. Over time, the further developments and services are becoming more complicated without a common, holistic frame for content and implementation. When gathering a large number of independent implementations, we experienced an increasing heterogeneity in content development but also in implementations of computing components.

This background is the major motivation for the development of an advanced framework based on long-term Knowledge Resources and integrated application components providing a valuable means of tackling the challenges. Nevertheless, in complex cases even major component groups cannot protect long-term challenges, if there is no basic framework architecture enabling to care for knowledge and computational implementation. The practice of creating solutions, which have to deal with the complements of knowledge suggests that flexible but nevertheless methodological, systematical approaches are required. The goal of this research is to create such knowledge-centric architecture, based on a wide range of multi-disciplinary implementations and practical case studies in different disciplines and dealing with different foci, for many years. While further developing and updating the knowledge related attributes, data, implementations, and solutions, all of them had to be revisited over time, improving and where necessary recreating implementation and developing content.

Knowledge resources and originary resources cover the complements of factual, conceptual, procedural, and metacognitive complements, e.g., from collections and referenced resources. The architecture presented here aims to seamlessly integrating separate roles of contributing parties, e.g., scientific staff creating research data, professional classification by experienced research library specialists, and developers of application components as well as services. The guideline was enabling to retain the knowledge required to resemble the inherent complexity of realia situations, real and material instead of abstract situations, while allowing lex parsimoniae principles of William of Ockham for problem solving.

Overall outcome and new insight gained from the practical knowledge resource object development and mining solutions based on the designed and implemented knowledge-centric architecture are presented in the following sections.

III. Fundaments and Previous Work

With one of the best and most solid works, Aristotle outlined the fundaments of terminology and of understanding knowledge [2] being an essential part of ‘Ethics’ [3]. Information sciences can very much benefit from Aristotle’s fundaments and a knowledge-centric approach, e.g., by Anderson and Krathwohl [4], but for building holistic and sustainable solutions they need to go beyond the available technology-based approaches and hypothesis [5] as analysed in Platons’ Phaidon. So far, there is no other practical advanced knowledge-centric architectural specification known, which implements these fundaments. Making a distinction and creating interfaces between methods and applications [6], the principles are based on the methodology of knowledge mapping [7]. The implementation can make use of objects and conceptual knowledge [8] and shows being able to build a base for application scenarios like associative processing [9] and advanced knowledge discovery [10]. Based on this background, during the last decades, a number of different case solutions were created, implemented, and realised on this fundament, including: Dynamical visualisation, knowledge mining, knowledge mapping, Content Factor, phonetic algorithms, Geoscientific Information Systems (GIS), Environmental Information Systems (EIS), cartographic mapping, service design, service management, and High End Computation. All such implementations include extensive use of LX Knowledge Resources, as explicitly representing the developed resources in this notation, and computation algorithms. This paper, presenting the new architecture, does not allow to illustrate too many implementation details. Therefore, an excerpt of practical solutions is cited, which have been reimplemented by the collaboration of the participated research groups and published, creating a base for this architecture. Representative examples are a) integrated systems and supercomputing resources used with phonetic algorithms and pattern matching [11] for knowledge mining [12], b) multi-dimensional context creation based on the methodology of Knowledge Mapping [13], and c) an exemplary resulting, widely used conceptual knowledge subset for geo-spatial scenarios [14]. The LX Knowledge Resources cover the factual, conceptual, procedural, and metacognitive complements in all cases, e.g., from collections and referenced resources.

An understanding of the essence and complexity of universal, multi-disciplinary knowledge can be achieved by taking a closer look on classification. The state-of-the-art of classifying ‘universal knowledge’ is the Universal Decimal Classification (UDC) [15] and its solid background, flexibility, and long history. The LX Knowledge Resources’ structure and the classification references [16] based on UDC [17], [18] are essential means for the processing workflows and evaluation. Both provide strong multi-disciplinary and multi-lingual support. For the research, all small unsorted excerpts of the Knowledge Resources objects only refer to main UDC-based classes, which for this publication are taken from the Multilingual Universal Decimal Classification Summary (UDCC Publication No. 088) [17] released by the UDC Consortium under the Creative Commons Attribution Share Alike 3.0 license [19] (first release 2009, subsequent update 2012). These components and their qualities are integrated in the resulting architecture with the methodologies and systematic use.

IV. Knowledge-centric Architecture

As discussed above, the presented results were achieved, based on the knowledge-centric architecture created for a series of previously implemented problem solutions and Knowledge Resources developments over the last years.

A. General Computation Architecture

The complements diagram of the implementation architecture [20][21][22] is shown in Figure 1. The major components are core resources and module resources. The result resources include object collections, which result from the application of core and module resources in arbitrary scenarios.

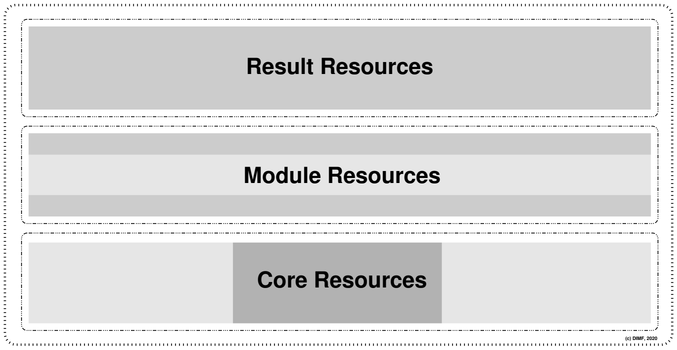


Figure 1. Complements diagram of the resources components architecture, including the three main complements of core, module, and result resources.

The sizes of this figure and the associated complements diagrams correspond, the following figures show complementary details from this context.

The core resources in this architecture comprise required resources. The complements diagram (Figure 2) shows the essential detail.

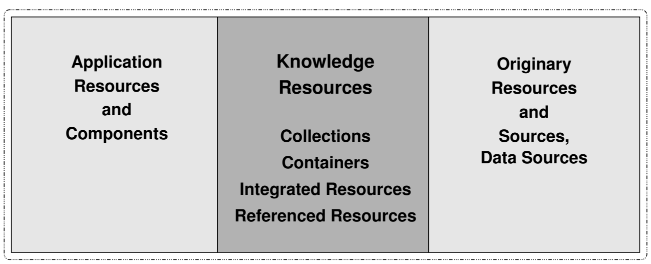


Figure 2. Computation architecture: Complements diagram of general core resources, from originary to knowledge and application resources.

The core resources can be divided into three categories: The central LX Knowledge Resources, originary resources, and application resources and components. The first, the Knowledge Resources, can include collections and containers as well as integrated resources and references to resources. The second, the originary resources, can include realia and original sources, which in many cases may have instances in the Knowledge Resources. The third, the application resources, can include implementations of algorithms, workflows, and procedures, which form applications and components. Instances of these components can also be employable in solutions due to their procedural nature, e.g., in module resources.

The complements diagram of general module resources is shown in Figure 3.

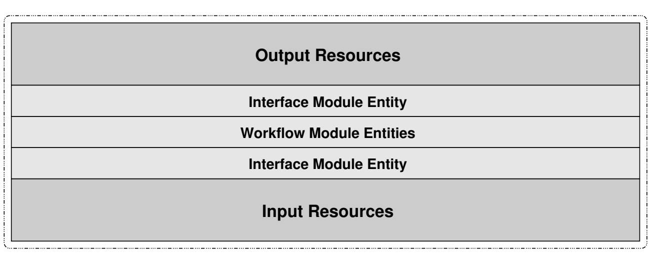


Figure 3. Computation architecture: Complements diagram of module resources, from input, interfaces and workflow entities to output.

A general set of module resources consists of input resources, modules, and output resources. The central workflow module entities are accompanied by interface module entities for input and output resources. For many architecture implementations, chains of module resources can be created, which can, for example, be used in pipeline and in parallel.

B. *Architecture Complements for Knowledge Mining*

For the case of knowledge mining, the complements diagram of the core resources is shown in Figure 4.

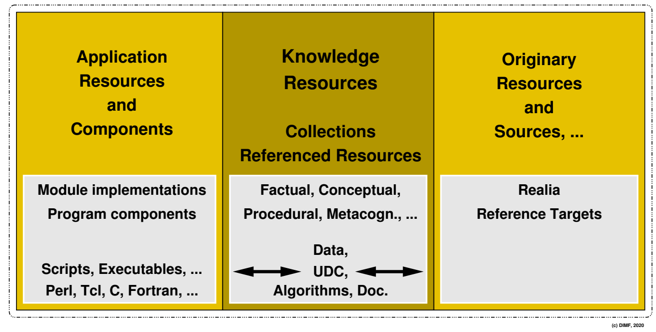


Figure 4. Knowledge Mining: Complements diagram of the core resources and examples of their contribution implementations.

Application resources and components are based on module implementations and program components for the knowledge mining realisation. Implementations employ scripting, high level languages, and third party components. Knowledge Resources and originary resources cover the complements of factual, conceptual, procedural, and metacognitive complements, e.g., from collections and references resources.

The respective complements diagram of a module resource for a text based knowledge mining implementation consists of several features (Figure 5).

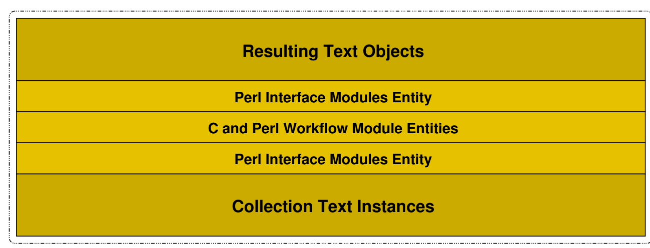


Figure 5. Knowledge Mining: Complements diagram of a module resource used for creating module chains for text based knowledge mining.

The input and output resources consist of text object instances in text based cases of knowledge mining. Here, the module entity implementations were implemented in C and Perl [23] for the respective implementations. The interface module entities are implemented in Perl, with the option to be on-the-fly generated within a workflow.

The knowledge-centric architecture does focus on resources and application scenarios, one of the most important is computation cases. Large computation workflow chains can be built with the architecture as was demonstrated with the reimplemented solutions for different cases, which were above referred to. An implementation sequence of module resources can be considered an intermediate step in building a workflow. Results within an implementation sequence can be considered intermediate results and instances, e.g., from the integrated mining of collection and container resources.

V. Object Development and Mining Cases

The goal was to create a knowledge-centric computation architecture, which allows a close integration of Knowledge Resources with wide spectra of complementary knowledge and flexible, efficient computational solutions, while being able to specify practically required roles for creation and long-term development. The knowledge-centric architecture can provide a base for an arbitrary range of use-cases and associated components. Two major groups of use-cases are

- resources creation and development and
- knowledge mining and selected associated methods.

A. Major Use-case Groups

Figure 6 shows an use-case diagram (UML, Unified Modeling Language) of the knowledge-centric computation architecture.

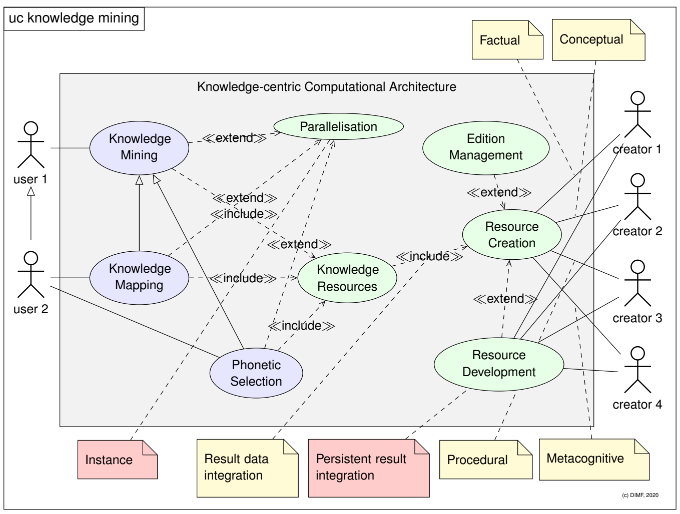


Figure 6. Use-case diagram of the knowledge-centric computation architecture: Two major use-case groups with four creator roles.

The excerpt illustrates an integrated view on the two groups of knowledge mining (bluish), which was implemented spanning knowledge mining use-cases and knowledge creation and development use-cases (greenish). In this widely deployed scenario, the implementation does have two main types of actors, namely creators and users. The use-cases have different actors, two 'user' roles and four 'creator' roles.

The selected system context is given by the grey box. The selected use-cases (ellipses) can be distinguished in use-cases for creators (greenish: resource creation, resource development, edition management, parallelisation) and use-cases for users (bluish: knowledge mining, knowledge mapping, phonetic selection).

Knowledge mining is supported by and using the cases of knowledge mapping and phonetic selection, which inherit to the knowledge mining instance the implemented methods and algorithms contributed by other user groups. For clearness, the creator and other roles for these two cases are not included in this diagram. Knowledge mining, mapping, and phonetic selection include the use-case of LX Knowledge Resources. The cases of this group are extended by parallelisation, respective computation, here instance based workflow parallelisation, which enables the computation-relevant optimisation for individual implementations and infrastructures.

The knowledge resources include the use-case of resource creation, which allows to integrate persistently added results. The use-case of resource creation itself is extended by the use-cases of resource development and edition management, which allows to define editions of resources for consistency in advanced complex application scenarios.

The use-case scenario reflects the professional practice of having separate roles for creating and developing factual, conceptual, procedural, and documentation, respective metacognitive knowledge. In most cases, different specialists are employed for creating and developing

- factual knowledge, e.g., research data and its documentation,
- conceptual knowledge, e.g., classification of knowledge objects,
- procedural knowledge, e.g., procedures, workflows, programs, and their respective documentation,
- metacognitive knowledge, e.g., documentation of experiences.

In practice, the creators are commonly represented by different groups, e.g., scientists, classification experts in scientific libraries, and designers of scientific algorithms and workflows.

B. Main Components

The core components of a basic knowledge mining implementation with the LX Knowledge Resources, based on the knowledge-centric computation architecture, can be summarised with a block diagram (Figure 7).

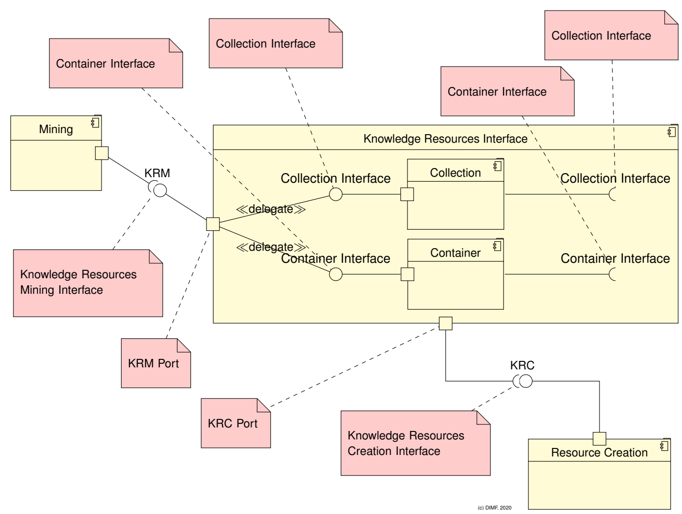


Figure 7. Block diagram of respective knowledge-centric computation architecture components: LX Knowledge Resources and interfaces excerpt.

The block diagram shows the LX Knowledge Resources and two types of knowledge object groups, namely object collections and containers. Each type and implementation can have individual and specialised interfaces. Knowledge mining is provided by an interface with the LX Knowledge Resources. The diagram also contains the interface block, due to the importance of the resource creation use-case. The individual groups have ports, interaction points, which can be used via interfaces, e.g., Knowledge Resources Creation (KRC) port and Knowledge Resources Mining (KRM) port. Components can have further interfaces, with and without delegating ports. It is common that independent resources are in many cases not necessarily orchestrated.

C. Activity Groups

A number of activities are associated with different components. An important activity regarding the resource creation is the creation-update (Figure 8).

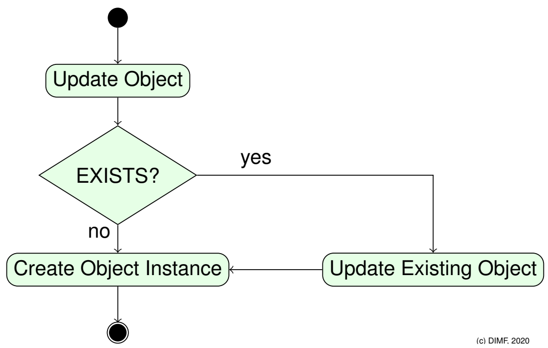


Figure 8. Activity diagram illustrating the essential object creation-update activities in the knowledge-centric computation architecture.

The resource creation component has to provide creation and update activities for the different creator groups. A simple but important example for resource creation and development is the creation of an object instance and respective updating an existing object with a new instance.

Start state is any state of the knowledge resources. End state is a new state of the knowledge resources. Regarding resource creation and development use-cases the start and end states should be considered intermediate states. As shown in the use-case diagram, professional practice affords the implementation of according activities for all required, specialised creator roles. A fundamental mining activity with knowledge resources is a resource request targeting to create an intermediate result (Figure 9).

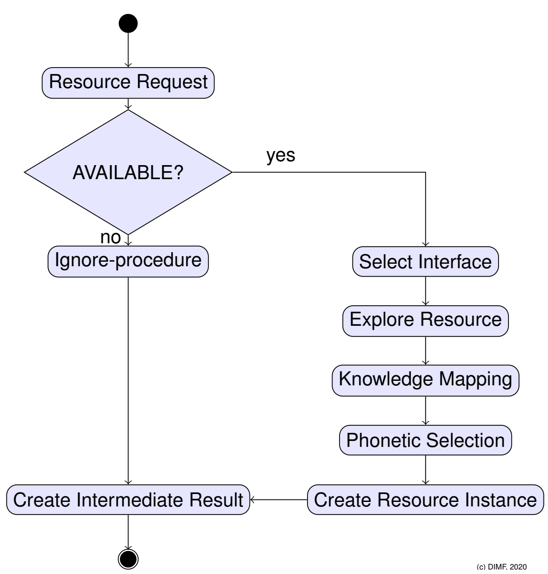


Figure 9. Activity diagram of a basic resource request for creating an intermediate knowledge mining result.

Start state is any state in a knowledge mining workflow chain. End state is a new state in a knowledge mining workflow chain. Regarding knowledge mining use-cases the start and end states should be considered intermediate states. If a resource is not available then an ignore-procedure continues for creating an intermediate result. The ignore-procedure can contribute its status to the workflow chain. If the resource is available then an interface is selected for the resource exploration. The exploration can use available activities, e.g., knowledge mapping and phonetic selection, in order to create a resource instance, which contributed to creating an intermediate result for the workflow chain.

Examples of activities are multi-dimensional context creation by knowledge mapping [7] and phonetic association, e.g., using Soundex [24][25][26]. Sample Soundex codes developed [12] are used for names in various textual, contextual, and linguistical situations, implemented in order to be integrated in a large number of situations.

VI. RESULTING OBJECT DEVELOPMENT ARCHITECTURE

Any sustainable long-term knowledge resources require an efficient concept for continuous content development. The fundamental concept for continuous content development implemented and practiced with this long-term research is shown in a Knowledge Resource Object Development (KROD) diagram (Figure 10).

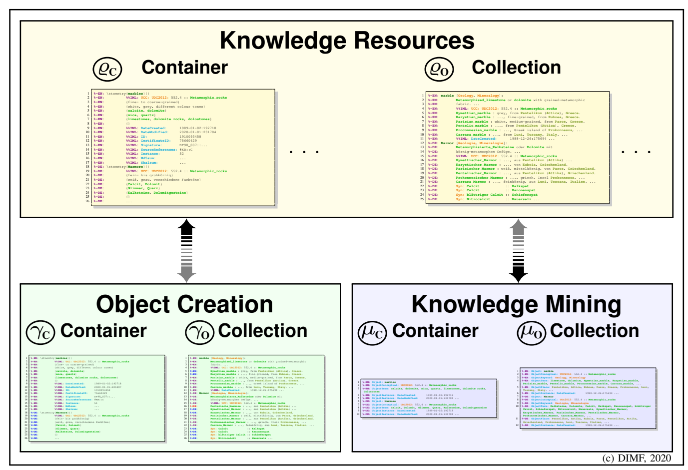


Figure 10. Knowledge Resource Object Development Diagram: Knowledge Resources, Object Creation, Knowledge Mining, containers and collections.

The diagram displays formalised activities associated with use-cases and their relations from the view of real object instances. The diagram illustrates activity groups and respective excerpts of objects.

Two major activities in context of LX Knowledge Resources (yellowish) are the creation of objects (greenish) and knowledge mining (bluish). Here, some major object groups are shown, collections and containers. The respective object instances (ϱ, γ, and μ groups) and properties are shown and further discussed in separate object instance representation views in the next sections. Different colours are used, which correspond to the discussed use-cases (Figures 6, 7, 8, and 9).

Table I shows the legend of colours/symbols used with the development components: The case groups (Knowledge Resources, KR; Object Creation, OC; Knowledge Mining, KM) and objects.

Table I. Legend for development components: Case groups (Knowledge Resources, KR; Object Creation, OC; Knowledge Mining, KM) and objects.

Development Component	*Colour Name*	*Colour*
Case group KR	(yellowish)	
Case group OC	(greenish)	
Case group KM	(bluish)	
Case group KR object	(darker yellowish)	
Case group OC object	(darker greenish)	
Case group KM object	(darker bluish)	
Objects, object terms	dark green	
Object terms' context	light green	
Object places' context	olive green	
Object keywords	orange	
Object instance data	cyan	
Object development-is	violet	
Object development-in	blue	
Object development-out	red	

The major groups of objects, container (C) objects and collection (O) objects [27], correspond with instances when used: Knowledge resources' objects in containers and collections, ϱ_{C} and ϱ_{O}, Object Creation's objects of containers and collections, γ_{C} and γ_{O}, and Knowledge Mining's objects from containers and collections, μ_{C} and μ_{O}.

A creation algorithm $\mathcal{C}$ realises a function $f_{\mathcal{C}}$ out of the set of input data γ and the set of output data ϱ, this is $f_{\mathcal{C}} : \gamma \rightarrow \varrho$.

A mining algorithm $\mathcal{M}$ realises a function $f_{\mathcal{M}}$ out of the set of input data ϱ and the set of output data μ, this is $f_{\mathcal{M}} : \varrho \rightarrow \mu$.

For the major C and O groups of objects, the algorithms for these two operations result in two groups of functions:

$$f_{\mathcal{C}} : \gamma_{\mathrm{C,O}} \longrightarrow \varrho_{\mathrm{C,O}} \tag{1}$$

$$f_{\mathcal{M}} : \varrho_{\mathrm{C,O}} \longrightarrow \mu_{\mathrm{C,O}} \tag{2}$$

Therefore, this architecture provides any algorithm a flexibility to be implemented by an arbitrary number of implementations, e.g., manual, automated, and hybrid approaches. Even more important is the development of the objects covering the complements of factual, conceptual, procedural, and metacognitive knowledge. Development policies and enforcement of policies are matter of implementing operation and services based on workflows and algorithms, which are beyond demonstration here. The following sections present the use-cases and object development, which are in focus of this paper.

VII. Object Development Case Results

The three types of instances of respective C and O objects, which are in focus of this research, are discussed in the following sections.

Excerpts of the LX Knowledge Resources provided for this research are shown using automatic component highlighting according to the above legend. Line-wise modification indicators were used for illustration in this paper.

A. Resources and Objects

The objects in the LX Knowledge Resources can contain content and documentation of factual, conceptual, procedural, and metacognitive knowledge. There are multi-fold means of attainment for this goal.

The selected resources contain more than 100,000 associated objects, entities, and references, especially referring to a potential of more than 4,000 minerals and mineral synonyms as well as possible 3,000,000 reference objects and conceptual knowledge in more than 50 languages.

The LX Knowledge Resources case group (yellowish), according to the KROD Diagram (Figure 10), is represented by excerpts illustrating object development examples from geoscientific container and collection objects. Figure 11 shows a small excerpt of a ϱ_{C} container object representation.

```
%-EN: \stoentry{marbles}{}
%-EN:           %%IML: UCC: UDC2012: 552.4 :: Metamorphic_rocks
%-EN:           {fine- to coarse-grained}
%-EN:           {white, grey, different colour tones}
%-EN:           {calcite, dolomite}
%-EN:           {mica, quartz}
%-EN:           {limestones, dolomite rocks, dolostones}
%-EN:           {}
%-EN:           %%IML: DateCreated:      1989-01-02:192718
%-EN:           %%IML: DateModified:     2020-01-01:231704
%-EN:           %%IML: ID:               1910093458
%-EN:           %%IML: CertificateID:    75660429
%-EN:           %%IML: Signature:        DF98_007::...
%-EN:           %%IML: SourceReferences: #KR::C
%-EN:           %%IML: Instance:         52
%-EN:           %%IML: Md5sum:           ...
%-EN:           %%IML: Sha1sum:          ...
%-DE: \stoentry{Marmore}{}
%-DE:           %%IML: UCC: UDC2012: 552.4 :: Metamorphic_rocks
%-DE:           {fein- bis grobkörnig}
%-DE:           {weiß, grau, verschiedene Farbtöne}
%-DE:           {Calcit, Dolomit}
%-DE:           {Glimmer, Quarz}
%-DE:           {Kalksteine, Dolomitgesteine}
%-DE:           {}
%-DE:           ...
```

Figure 11. Object instance representation: LX Knowledge Resources KR, container object (excerpt).

The auto-highlighting emphasises various knowledge in context of this object instance. The object entity of type stone is 'marbles' (dark green). Material components (light green) are calcite and dolomite (major components), mica and quartz (minor components). Educts are limestone, dolomite rocks etc. Conceptual knowledge is available from the Universal Classified Classification (UCC) concordances [10] as UDC2012 (orange), classifying the object with a reference to the major group `UDC:552.4`, metamorphic rocks. Here, e.g., the conceptual knowledge refers to metamorphic rocks. The object carries multi-lingual instances, English and German (indicated by `%-EN:` and `%-DE:`). This excerpt is instance number `52`. Here, the shown excerpts indicate that instances of this object entity were created and developed over significant date range. Figure 12 shows a small excerpt of a ϱ_{O} collection object representation.

```
%-EN: marble [Geology, Mineralogy]:
%-EN:         Metamorphised_limestone or dolomite with grained-metamorphic
%-EN:         fabric. ...
%-EN:         %%IML: UCC: UDC2012: 552.4 :: Metamorphic_rocks
%-EN:         Hymettian_marble : grey, from Pentelikon (Attica), Greece.
%-EN:         Karystian_marble : ..., fine-grained, from Euboea, Greece.
%-EN:         Parisian_marble : white, medium-grained, from Paros, Greece.
%-EN:         Pentelic_marble : ..., from Pentelikon (Attica), Greece.
%-EN:         Proconnesian_marble : ..., Greek island of Prokonnesos, ...
%-EN:         Carrara_marble : ..., from Luni, Tuscany, Italy. ...
%-EN:         %%IML: DateCreated:      1988-12-26:175694 ...
%-DE: Marmor [Geologie, Mineralogie]:
%-DE:         Metamorphisierte_Kalksteine oder Dolomite mit
%-DE:         körnig-metamorphem Gefüge. ...
%-DE:         %%IML: UCC: UDC2012: 552.4 :: Metamorphic_rocks
%-DE:         Hymettischer_Marmor : ..., aus Pentelikon (Attika) ...
%-DE:         Karystischer_Marmor : ..., von Euboia, Griechenland.
%-DE:         Parisischer_Marmor : weiß, mittelkörnig, von Paros, Griechenland.
%-DE:         Pentelischer_Marmor : ..., aus Pentelikon (Attika), Griechenland.
%-DE:         Prokonnesischer_Marmor : ..., griech. Insel Prokonnesos, ...
%-DE:         Carrara_Marmor : ..., feinkörnig, aus Luni, Toscana, Italien. ...
%-DE:         Syn: Calcit                :: Kalkspat
%-DE:         Syn: Calcit                :: Kanonenspat
%-DE:         Syn: blättriger Calcit :: Schieferspat
%-DE:         Syn: Nitrocalcit           :: Mauersalz ...
```

Figure 12. Object instance representation: LX Knowledge Resources KR, collection object (excerpt).

The LX Knowledge Resources' collection object contains a 'marble' and 'Marmor' entity (dark green), material components (light green), keywords ('Geology', 'Mineralogy'), conceptual knowledge from Universal Classified Classification (UCC) concordance references as UDC2012 (orange), classifying the object with a reference to the major group `UDC:552.4`, metamorphic rocks.

The representations and highlighting also show the lingual contexts (violet), e.g., English, dates regarding the object instance and so on. The collection object carries multi-lingual instances, English and German (indicated by `%-EN:` and `%-DE:`).

The object instances carry their instance data (cyan). Instance representation, e.g., identification, is available from the object and entity instances, Object Envelopes (OEN), [28]. A larger excerpt of the instance data of the ϱ_{o} object is shown in Figure 13.

```
%-EN:        %%IML: DateCreated:       1988-12-26:175694
%-EN:        %%IML: DateModified:      2020-01-01:232927
%-EN:        %%IML: ID:                1909042178
%-EN:        %%IML: CertificateID:     74581231
%-EN:        %%IML: Signature:         DF98_007::...
%-EN:        %%IML: SourceReferences: #KR::O
%-EN:        %%IML: Instance:          119
%-EN:        %%IML: Md5sum:            ...
%-EN:        %%IML: Sha1sum:           ...
```

Figure 13. Object instance representation: Instance data excerpt of LX Knowledge Resources collection object (Figure 12).

The excerpt shows creation and modification dates, ids, references, and check sums.

B. Object Creation and Development

The object creation case group (greenish), according to the KROD Diagram (Figure 10), is represented by following the developments of object examples presented in the above resources section. Figure 14 shows a small excerpt of a γ_{C} container object representation.

```
%-EN: \stoentry{marbles}{}
%-EN:          %%IML: UCC: UDC2012: 552.4 :: Metamorphic_rocks
%-EN:          {fine- to coarse-grained}
%-EN:          {white, grey, different colour tones}
%-EN:          {calcite, dolomite}
%-EN:          {mica, quartz}
%-EN:          {limestones, dolomite rocks, dolostones}
%-EN:          {}
%-EN:          %%IML: DateCreated:       1989-01-02:192718
%-EN:          %%IML: DateModified:      2020-01-01:225807
%-EN:          %%IML: ID:                1910093458
%-EN:          %%IML: CertificateID:     75660429
%-EN:          %%IML: Signature:         DF98_007::...
%-EN:          %%IML: SourceReferences: #KR::C
%-EN:          %%IML: Instance:          51
%-EN:          %%IML: Md5sum:            ...
%-EN:          %%IML: Sha1sum:           ...
%>DE: \stoentry{Marmore}{}
%>DE:          %%IML: UCC: UDC2012: 552.4 :: Metamorphic_rocks
%>DE:          {fein- bis grobkörnig}
%>DE:          {weiß, grau, verschiedene Farbtöne}
%>DE:          {Calcit, Dolomit}
%>DE:          {Glimmer, Quarz}
%>DE:          {Kalksteine, Dolomitgesteine}
%>DE:          {}
%>DE:          ...
```

Figure 14. Object instance representation: Object Creation OC container object (excerpt).

The OC object illustrates the object creation and development process, an update of an existing object on base of a previously created object instance in the Knowledge Resources. This excerpt is based on instance number `51`, creating instance number `52` (Figure 11). The excerpt shows the addition of a language entry to the container. Here, an entity in German language is added to the object. In this case, the added material is indicated by line-based blue markers, `%>DE:`. Figure 15 shows a small excerpt of a γ_{O} collection object representation.

```
%-EN: marble [Geology, Mineralogy]:
%-EN:        Metamorphised_limestone or dolomite with grained-metamorphic
%-EN:        fabric. ...
%-EN:        %%IML: UCC: UDC2012: 552.4 :: Metamorphic_rocks
%>EN:        Hymettian_marble : grey, from Pentelikon (Attica), Greece.
%>EN:        Karystian_marble : ..., fine-grained, from Euboea, Greece.
%-EN:        Parisian_marble : white, medium-grained, from Paros, Greece.
%-EN:        Pentelic_marble : ..., from Pentelikon (Attica), Greece.
%-EN:        Proconnesian_marble : ..., Greek island of Prokonnesos, ...
%-EN:        Carrara_marble : ..., from Luni, Tuscany, Italy. ...
%-EN:        %%IML: DateCreated:      1988-12-26:175694 ...
%-DE: Marmor [Geologie, Mineralogie]:
%-DE:        Metamorphisierte_Kalksteine oder Dolomite mit
%-DE:        körnig-metamorphem Gefüge. ...
%-DE:        %%IML: UCC: UDC2012: 552.4 :: Metamorphic_rocks
%<DE:        Pentelische_Marmore : ..., aus Pentelikon (Attika) ...
%>DE:        Hymettischer_Marmor : ..., aus Pentelikon (Attika) ...
%>DE:        Karystischer_Marmor : ..., von Euboia, Griechenland.
%-DE:        Parisischer_Marmor : weiß, mittelkörnig, von Paros, Griechenland.
%-DE:        Pentelischer_Marmor : ..., aus Pentelikon (Attika), Griechenland.
%-DE:        Prokonnesischer_Marmor : ..., griech. Insel Prokonnesos, ...
%-DE:        Carrara_Marmor : ..., feinkörnig, aus Luni, Toscana, Italien. ...
%>DE:        Syn: Calcit            :: Kalkspat
%>DE:        Syn: Calcit            :: Kanonenspat
%>DE:        Syn: blättriger Calcit :: Schieferspat
%>DE:        Syn: Nitrocalcit       :: Mauersalz ...
```

Figure 15. Object instance representation: Object Creation OC collection object (excerpt).

The excerpt shows modification of entries in a collection object. Entries are added (`%>EN:` and `%>DE:`, blue colour) and removed (`%<DE:`, red colour). The object creation adds new types of rock in two languages and a number of synonyms, which are unique to German language. A larger excerpt of the instance data of the γ_{O} object is shown in Figure 16.

```
%-EN:        %%IML: DateCreated:       1988-12-26:175694
%<EN:        %%IML: DateModified:      2017-05-30:221532
%>EN:        %%IML: DateModified:      2020-01-01:231704
%-EN:        %%IML: ID:                1909042178
%-EN:        %%IML: CertificateID:     74581231
%-EN:        %%IML: Signature:         DF98_007::...
%-EN:        %%IML: SourceReferences: #KR::O
%<EN:        %%IML: Instance:          118
%>EN:        %%IML: Instance:          119
%<EN:        %%IML: Md5sum:            ...
%>EN:        %%IML: Md5sum:            ...
%<EN:        %%IML: Sha1sum:           ...
%>EN:        %%IML: Sha1sum:           ...
```

Figure 16. Object instance representation: Instance data excerpt of Object Creation collection object (Figure 15).

The excerpt in state of development contains references of instances `118` and `119` of the ϱ_{O} object.

VIII. Knowledge Mining Case Results

A. Objects and mining

The Knowledge Mining case group (bluish), according to the KROD Diagram (Figure 10), is represented by following the developments of object examples presented in the above resources and creation/development sections. Figure 17 shows a small excerpt of a μ_{C} container object extract representation.

```
%-EN: Object: marbles
%-EN: ObjectConceptual: UDC2012: 552.4 :: Metamorphic_rocks
%-EN: ObjectTerm: calcite, dolomite, mica, quartz, limestones, dolomite rocks, dolostones
%-EN: ObjectInstance: DateCreated:       1989-01-02:192718
%-EN: ObjectInstance: DateModified:      2020-01-01:231704 ...
%-DE: Object: Marmore
%-DE: ObjectConceptual: UDC2012: 552.4 :: Metamorphic_rocks
%-DE: ObjectTerm: Calcit, Dolomit, Glimmer, Quarz, Kalksteine, Dolomitgesteine
%-DE: ObjectInstance: DateCreated:       1989-01-02:192718
%-DE: ObjectInstance: DateModified:      2020-01-01:231704 ...
```

Figure 17. Object instance representation: Knowledge Mining KM container object extract (excerpt).

Figure 18 shows a small excerpt of an extracted μ_{O} collection object representation.

```
%-EN: Object: marble
%-EN: ObjectConceptual: UDC2012: 552.4 :: Metamorphic_rocks
%-EN: ObjectKeyword: Geology, Mineralogy
%-EN: ObjectTerm: limestone, dolomite, Hymettian_marble, Karystian_marble,
 Parisian_marble, Pentelic_marble, Proconnesian_marble, Carrara_marble, ...
%-EN: ObjectPlace: Pentelikon, Attica, Euboea, Paros, Greece, Prokonnesos, Luni,
  Tuscany, Italy. ...
%-EN: ObjectInstance: DateCreated:      1988-12-26:175694 ...
%-DE: Object: Marmor
%-DE: ObjectConceptual: UDC2012: 552.4 :: Metamorphic_rocks
%-DE: ObjectKeyword: Geologie, Mineralogie
%-DE: ObjectTerm: Kalksteine, Dolomite, Calcit, Kalkspat, Kanonenspat,
 blättriger Calcit, Schieferspat, Nitrocalcit, Mauersalz, Hymettischer_Marmor,
 Karystischer_Marmor, Parisischer_Marmor, Pentelischer_Marmor,
 Prokonnesischer_Marmor, Carrara_Marmor, ...
%-DE: ObjectPlace: Pentelikon, Attika, Euboia, Paros, Pentelikon, Attika,
 Griechenland, Prokonnesos, Luni, Toscana, Italien, ...
%-DE: ObjectInstance: DateCreated:      1988-12-26:175694 ...
```

Figure 18. Object instance representation: Knowledge Mining KM collection object extract (excerpt).

A larger excerpt of the instance data of the extract $\mu_{\rm O}$ object is shown in Figure 19.

```
%-EN:        %%IML: DateCreated:      1988-12-26:175694
%-EN:        %%IML: DateModified:     2020-01-01:232927
%-EN:        %%IML: ID:               1909042178
%-EN:        %%IML: CertificateID:    74581231
%-EN:        %%IML: Signature:        DF98_007::...
%-EN:        %%IML: SourceReferences: #KR::O
%-EN:        %%IML: Instance:         119 [extract]
%-EN:        %%IML: Md5sum:           ...
%-EN:        %%IML: Sha1sum:          ...
```

Figure 19. Object instance representation: Instance data excerpt of Knowledge Mining collection object extract (Figure 18).

The purpose of intermediate results in a universal context is to provide a high flexibility and modularity in workflows. It must be possible to create and re-use algorithms, which are taking care for arbitrary scenarios. An integrated intermediate result of such case, a resource request on 'marble', creating an intermediate result including object instances from $\mu_{\rm O}$ and $\mu_{\rm C}$, is displayed in Figure 20.

```
%-EN: Object: marble, marbles
%-EN: ObjectConceptual: UDC2012: 552.4 :: Metamorphic_rocks
%-EN: ObjectKeyword: Geology, Mineralogy
%-EN: ObjectTerm: limestone, dolomite, calcite, mica, quartz, limestones,
 dolomite rocks, dolostones, Hymettian_marble, Karystian_marble, Parisian_marble
 , Pentelic_marble, Proconnesian_marble, Carrara_marble, ...
%-EN: ObjectPlace: Pentelikon, Attica, Euboea, Paros, Greece, Prokonnesos, Luni,
  Tuscany, Italy, ...
%-EN: ObjectInstance: SourceReferences: #KR::C::1910093458::52 #
 KR::O::1909042178::119
%-DE: Object: Marmor, Marmore
%-DE: ObjectConceptual: UDC2012: 552.4 :: Metamorphic_rocks
%-DE: ObjectKeyword: Geologie, Mineralogie
%-DE: ObjectTerm: Kalksteine, Dolomite, Dolomit, Calcit, Glimmer, Quarz,
 Kalksteine, Dolomitgesteine, Kalkspat, Kanonenspat, blättriger Calcit,
 Schieferspat, Nitrocalcit, Mauersalz, Hymettischer_Marmor, Karystischer_Marmor,
  Parisischer_Marmor, Pentelischer_Marmor, Prokonnesischer_Marmor,
 Carrara_Marmor, ...
%-DE: ObjectPlace: Pentelikon, Attika, Euboia, Paros, Griechenland, Prokonnesos,
  Luni, Toscana, Italien, ...
%-DE: ObjectInstance: SourceReferences: #KR::C::1910093458::52 #
 KR::O::1909042178::119
```

Figure 20. Object instance representation: Knowledge Mining intermediate result from objects' extracts (excerpt).

This excerpt is only one of many possible, different intermediate results. The algorithms of creating intermediate results allow arbitrary integration and formalisation, abstraction, reduction. Any available methods can contribute, including mining algorithms, knowledge mapping, and phonetic selection. This example of an integrated intermediate result contains structured information from multi-lingual object entities, material, and location information, from both collection and container objects. The intermediate result resulting from the mining process is integrating knowledge from both multi-lingual collection and container instances (`#KR::C::1910093458::52` and `#KR::O::1909042178::119`), as referring to the previously shown instances `52` of the $\varrho_{\rm C}$ object and `119` of the $\varrho_{\rm O}$ object. Here, the excerpt shows references from a stones container and a geoscientific collection. Further width and depth of integrated knowledge in this context can include any multi-disciplinary references, e.g., references can range from mineralogy containers to geoscientific synonym containers.

IX. Discussion

The continuous object development and knowledge mining cases results were efficiently and effectively supported deploying the features provided by the object development architecture and functions.

- Development of resources and objects were successfully implemented and demonstrated for object groups in collections and containers.
- Object structures have shown to allow highest flexibility. All components support multi-disciplinary knowledge objects. All groups and objects allow flexible object and context references, support of knowledge complements, e.g., conceptual knowledge via UDC, and support multi-lingual content and access.
- All components can ensure consistency of resources, object re-use, advanced knowledge mining, creation and deployment of algorithms, and object integration.
- The fundamental, universal framework is already successfully developed and used in practice for many years before publication.

The knowledge-centric architecture proved being a flexible, reliable, and robust fundament for object creation and development as well as for advanced knowledge mining, managing and using LX Knowledge Resources objects' collections and containers for many years. The architecture provides the flexibility that workflow chain modules and whole workflow chains can be employed sequential or parallel. The components in general are not limited by the architecture to be implemented for synchronous or asynchronous accesses if required for arbitrary algorithms and workflows.

The implementations for practical case studies built upon this architecture span different disciplines and deal with different foci. The excerpted cases include general, simplified cases of knowledge mining and practical knowledge development scenarios from realisations, which were implemented for large practical solutions. These cases are relevant because of the professional background and practice required to deal with development of resources and application components for long-term tasks.

In complex scenarios, different disciplines contribute fulfilling different tasks. In case of knowledge creation and development and its valorisation different specialised expertise is required. In general, content and applications are created by different disciplines. Even different aspects of content may require different specialists groups, different roles, e.g., natural sciences research data and conceptual valorisation are often done by scientists from a respective discipline and information scientists. Many components have to be revisited and improved over time as the results and facilities should be continued and preserved and be available for long-term. In the implementation cases, factual, conceptual, procedural, and metacognitive knowledge is cared for by different experts. The architecture allows flexible and efficient separation of roles. For example, research data can be created by a role and can at any time be amended with classification and procedural documentation by experienced research library specialist and researcher roles.

The LX Knowledge Resources are containing a lot more content and references than can be used at present time in most cases. The architecture allowed to support retaining the associated knowledge required to resemble the intrinsic complexity of realia situations while implementing selected solutions for isolated as well as complex situations. The development of knowledge mining and the provisioning of services based on these tasks can continuously be done by application developers, accessing the continuously extendable LX Knowledge Resources.

X. Conclusion

This paper presented a knowledge resources' view on object development based on a new knowledge-centric architecture. The paper presented the architecture and discussed two successfully implemented major use-cases, object development and knowledge mining. The implementation supports all the features, which were targeted and should be provided regarded the architecture and object development.

Knowledge resource object development and respective object instance representation for resources, object creation and development, and knowledge mining were presented for different types of object groups, namely collections and containers.

The paper presented consistent view on activities and real object instances from use-case scenarios in context with the research results of creating a knowledge-centric computation architecture. The resulting architecture was developed in continuous cross development of multi-disciplinary, multilingual knowledge resources and practical knowledge-centric solutions. This paper presented the major qualities of the computation architecture. The practical employment of the architecture was illustrated for advanced knowledge mining and practical development uses-cases. The contributing research collaboration achieved to create a practical approach for a knowledge-centric computation architecture, which allows the methodological and systematical development and employment of components, including knowledge resources. The architecture covers the creation of flexible solutions, which allow to most widely employ the complements of knowledge.

The long-term knowledge resource scenarios and the knowledge-centric architecture and use-cases proved in practice to support the object development, the seamless separation and integration of roles for different disciplines and tasks while implementing and realising solutions based on knowledge resources and computation. In addition to the implemented and referred case studies, it was shown that object development and major use-cases can be efficiently managed. Especially, on the one hand, knowledge creation and development can be professionally dealt with by groups from responsible disciplines. On the other hand, knowledge mining can be based on the work of these disciplines while service based use and implementation can be given different roles, relying on the resources being continuously in development.

Future research will concentrate on further extending and developing knowledge resources in order to foster the creation of content bases and to provide long-term capacities and creating new advanced algorithms and mining workflows for enabling fundaments for new insight, participating different institutions and roles, based on the knowledge-centric computation architecture and knowledge resources.

Acknowledgment

We are grateful to the "Knowledge in Motion" (KiM) long-term project, Unabhängiges Deutsches Institut für Multidisziplinäre Forschung (DIMF), for partially funding implementations, case studies, publications, and for granting access to LX Knowledge Resources under grants D2016F5P04648, D2017F1P04812, and D2018F6P04938 and to its senior scientific members and members of the permanent commission of the science council and the board of trustees, especially to Dr. Friedrich Hülsmann, Gottfried Wilhelm Leibniz Bibliothek (GWLB) Hannover, to Dipl.-Biol. Birgit Gersbeck-Schierholz, Leibniz Universität Hannover, to Dipl.-Ing. Martin Hofmeister, Hannover, and to Olaf Lau, Hannover, Germany, for collaboration, practical multi-disciplinary case studies, and the analysis and implementation of advanced concepts. We are grateful to Dipl.-Ing. Hans-Günther Müller, HPE, Germany, for his excellent contributions and assistance, providing practical private cloud and storage solutions. We are grateful to all national and international partners in the Geo Exploration and Information cooperations for their constructive and transdisciplinary support. We are grateful to the Science and High Performance Supercomputing Centre (SHPSC) for long-term support and The International ARS Science and History Network for providing multi-disciplinary application scenarios and assistance. / DIMF-PIID-DF98_007.

References

[1] C.-P. Rückemann, "A Knowledge-centric Computation Architecture and the Case of Knowledge Mining," in Proceedings of The Ninth International Conference on Advanced Communications and Computation (INFOCOMP 2019), July 28 – August 1, 2019, Nice, France. XPS Press, Wilmington, Delaware, USA, 2019, pp. 35–40, ISSN: 2308-3484, ISBN: 978-1-61208-732-0, URL: http://www.thinkmind.org/index.php?view=article&articleid=infocomp_2019_3_20_60044 [accessed: 2020-05-02].

[2] Aristotle, Nicomachean Ethics, 2008, (Written 350 B.C.E.), Translated by W. D. Ross, Provided by The Internet Classics Archive, URL: http://classics.mit.edu/Aristotle/nicomachaen.html [accessed: 2020-05-02].

[3] Aristotle, The Ethics of Aristotle, 2005, Project Gutenberg, eBook, EBook-No.: 8438, Release Date: July, 2005, Digitised Version of the Original Publication, Produced by Ted Garvin, David Widger, and the DP Team, Edition 10, URL: http://www.gutenberg.org/ebooks/8438 [accessed: 2020-05-02].

[4] L. W. Anderson and D. R. Krathwohl, Eds., A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Allyn & Bacon, Boston, MA (Pearson Education Group), USA, 2001, ISBN-13: 978-0801319037.

[5] Plato, Phaedo, 2008, (Written 360 B.C.E.), Translated by Benjamin Jowett, Provided by The Internet Classics Archive, URL: http://classics.mit.edu/Plato/phaedo.html [accessed: 2020-05-02].

[6] C.-P. Rückemann and F. Hülsmann, "Significant Differences: Methodologies and Applications," KiMrise, Knowledge in Motion Meeting, November 27, 2017, Knowledge in Motion, Hannover, Germany, 2017.

[7] C.-P. Rückemann, "Methodology of Knowledge Mapping for Arbitrary Objects and Entities: Knowledge Mining and Spatial Representations – Objects in Multi-dimensional Context," in Proceedings of The Tenth International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2018), March 25 – 29, 2018, Rome, Italy. XPS Press, Wilmington, Delaware, USA, 2018, pp. 40–45, ISSN: 2308-393X, ISBN-13: 978-1-61208-617-0, URL: http://www.thinkmind.org/index.php?view=article&articleid=geoprocessing_2018_3_20_30078 [accessed: 2020-05-02].

[8] C.-P. Rückemann, "Creation of Objects and Concordances for Knowledge Processing and Advanced Computing," in Proceedings of The Fifth International Conference on Advanced Communications and Computation (INFOCOMP 2015), June 21–26, 2015, Brussels, Belgium. XPS Press, 2015, ISSN: 2308-3484, ISBN-13: 978-1-61208-416-9, URL: http://www.thinkmind.org/index.php?view=article&articleid=infocomp_2015_4_30_60038 [accessed: 2020-05-02].

[9] C.-P. Rückemann, "Advanced Association Processing and Computation Facilities for Geoscientific and Archaeological Knowledge Resources Components," in Proceedings of The Eighth International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2016), April 24 – 28, 2016, Venice, Italy. XPS Press, 2016, pp. 69–75, ISSN: 2308-393X, ISBN-13: 978-1-61208-469-5, URL: http://www.thinkmind.org/download.php?articleid=geoprocessing_2016_4_20_30144 [accessed: 2020-05-02].

[10] C.-P. Rückemann, "Advanced Knowledge Discovery and Computing based on Knowledge Resources, Concordances, and Classification," International Journal On Advances in Intelligent Systems, vol. 9, no. 1&2, 2016, pp. 27–40, ISSN: 1942-2679, URL: http://www.iariajournals.org/intelligent_systems/intsys_v9_n12_2016_paged.pdf [accessed: 2020-05-02].

[11] "Perl Compatible Regular Expressions (PCRE)," 2020, URL: https://www.pcre.org/ [accessed: 2020-05-02].

[12] C.-P. Rückemann, "Archaeological and Geoscientific Objects used with Integrated Systems and Scientific Supercomputing Resources," International Journal on Advances in Systems and Measurements, vol. 6, no. 1&2, 2013, pp. 200–213, ISSN: 1942-261x, URL: http://www.iariajournals.org/systems_and_measurements/sysmea_v6_n12_2013_paged.pdf [accessed: 2020-05-02].

[13] C.-P. Rückemann, "Multi-dimensional Context Creation Based on the Methodology of Knowledge Mapping," International Journal on Advances in Software, vol. 11, no. 3&4, 2018, pp. 286–298, ISSN: 1942-2628, URL: http://www.iariajournals.org/software/soft_v11_n34_2018_paged.pdf [accessed: 2020-05-02].

[14] C.-P. Rückemann, "Superordinate Knowledge Based Comprehensive Subset of Conceptual Knowledge for Practical Geo-spatial Application Scenarios," in Proceedings of The Eleventh International Conference on Advanced Geographic Information Systems, Applications, and Services (GEOProcessing 2019), February 24 – 28, 2019, Athens, Greece. XPS Press, Wilmington, Delaware, USA, 2019, pp. 52–58, ISSN: 2308-393X, ISBN: 978-1-61208-687-3, URL: http://www.thinkmind.org/index.php?view=article&articleid=geoprocessing_2019_3_30_30039 [accessed: 2020-05-02].

[15] UDC, Universal Decimal Classification. British Standards Institute (BSI), 2005, complete Edition, ISBN: 0-580-45482-7, Vol. 1 and 2.

[16] C.-P. Rückemann, "Enabling Dynamical Use of Integrated Systems and Scientific Supercomputing Resources for Archaeological Information Systems," in Proceedings of the International Conference on Advanced Communications and Computation (INFOCOMP 2012), October 21–26, 2012, Venice, Italy, 2012, pp. 36–41, ISBN: 978-1-61208-226-4, URL: http://www.thinkmind.org/download.php?articleid=infocomp_2012_3_10_10012 [accessed: 2020-05-02].

[17] "Multilingual Universal Decimal Classification Summary," 2012, UDC Consortium, 2012, Web resource, v. 1.1. The Hague: UDC Consortium (UDCC Publication No. 088), URL: http://www.udcc.org/udcsummary/php/index.php [accessed: 2020-05-02].

[18] "UDC Online," 2020, URL: http://www.udc-hub.com [accessed: 2020-05-02].

[19] "Creative Commons Attribution Share Alike 3.0 license," 2012, URL: http://creativecommons.org/licenses/by-sa/3.0/ [accessed: 2020-05-02].

[20] B. Oestereich and A. Scheithauer, Analyse und Design mit der UML 2.5: Objektorientierte Softwareentwicklung (in English: Analysis and Design With the UML 2.5: Object-oriented Software Development), 11th ed. De Gruyter Oldenbourg, 2013, ISBN: 978-3486721409.

[21] "Unified Modeling Language (UML)," 2020, URL: http://uml.org/ [accessed: 2020-05-02].

[22] "About the Unified Modeling Language Specification 2.5.1," Dec. 2017, OMG – Object Management Group, URL: https://www.omg.org/spec/UML/ [accessed: 2020-05-02].

[23] "The Perl Programming Language," 2020, URL: https://www.perl.org/ [accessed: 2020-05-02].

[24] R. C. Russel and M. K. Odell, "U.S. patent 1261167," 1918, (Soundex algorithm), patent issued 1918-04-02.

[25] D. E. Knuth, The Art of Computer Programming: Sorting and Searching. Addison-Wesley, 1973, vol. 3, ISBN: 978-0-201-03803-3, OCLC: 39472999, pp. 391–392.

[26] National Archives and Records Administration (NARA), "The Soundex Indexing System," 2007, 2007-05-30, URL: http://www.archives.gov/research/census/soundex.html [accessed: 2020-05-02].

[27] C.-P. Rückemann, "Integration of Knowledge Resources and Advanced Association Processing for Geosciences and Archaeology," International Journal on Advances in Systems and Measurements, vol. 9, no. 3&4, 2016, pp. 485–495 ISSN: 1942-261x, URL: http://www.iariajournals.org/systems_and_measurements/sysmea_v9_n34_2016_paged.pdf [accessed: 2020-05-02].

[28] C.-P. Rückemann, "High End Computing Using Advanced Archaeology and Geoscience Objects," International Journal On Advances in Intelligent Systems, vol. 6, no. 3&4, 2013, ISSN: 1942-2679, LCCN: 2008212456, URL: http://www.iariajournals.org/intelligent_systems/intsys_v6_n34_2013_paged.pdf [accessed: 2020-05-02].

A Feature Selection and Extraction Method from Time-Frequency Images

Dorel Aiordăchioaie
"Dunarea de Jos" University of Galati
Electronics and Telecom. Department
Galati, Romania
e-mail: Dorel.Aiordachioaie@ugal.ro

Theodor D. Popescu
National Institute for Research and Development in Informatics
Bucharest, Romania
e-mail: Theodor.Popescu@ici.ro

Bogdan Dumitrascu
"Dunarea de Jos" University of Galati
Electronics and Telecom. Department
Galati, Romania
e-mail: Bogdan.Dumitrascu@ugal.ro

***Abstract*—Time-frequency image processing is considered in the context of change detection and diagnosis purposes based on signal processing paradigm. A method for selection and extraction of features from time-frequency is considered and evaluated. New images are obtained by applying a criterion based on the contours generated by the main components of the analyzed time-frequency image. The transformed images are less complex and could be white and black only. Features based on statistical moments are considered, selected and used to define discriminant functions, in order to improve the results of the classification. The features include the number of the contours, the average area defined by the contours, the variance of the areas and the Renyi entropies. As case study, signals coming from vibration generated by faults in bearings are considered. The main output of the paper is the method of the feature selection and extraction from time-frequency images.**

Keywords - signal; image; time-frequency transform; signal processing; feature selection; classification.

I. INTRODUCTION

Incipient fault detection in mechanical processes by vibration monitoring is an important activity for various goals such as predictive maintenance, safety and product cost optimization. The paper presents a method for feature selection and extraction from time-frequency images of the vibration signals generated by various faults in the bearings of the rotating machines for classification purposes mainly and represents an extended and enhanced version of [1]. More facts and references are presented in the state of the art, the structure of the method is introduced, more details from the selection process of the features, and more results from experiments are presented also.

An important activity in industry, for safe work and quality of the products, is the Change Detection and Diagnosis (CDD) in various processes. These two activities are parts of a wider domain, called condition-based and predictive maintenance, as described in some excellent books with theory and applications [2] [3] [4]. In the field of vibrational processes, i.e., processes that generate mechanical vibrations, with or without faults or damages, advanced signal processing algorithms are intensively used to elaborate accurate and robust algorithms for process diagnosis [5] [6] [7].

One of the more complex signal processing method is based on time-frequency transform, and next on time-frequency images, as described in [8] [9] [10]. The structure of such processing chain is presented in Figure 1. Signals from the process under study are pre-processed both in continuous and discrete time, mainly by filtering and scaling. Next, a time sliding window is considered for the computation of the time-frequency transform.

The parameters of the sliding window depend on the statistical properties of the analyzed signals, to meet the condition of the statistical stationarity. The coefficients of the time-frequency transform are considered as elements of an image. From this point all processing steps are based on image processing, for various tasks, as fault detection and diagnosis. Finally, from methodological point of view a set of papers and practical examples are available as [11] [12].

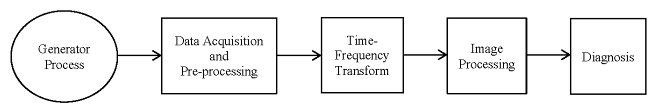


Figure 1. The block structure of signal processing for CDD

This work considers the last block before diagnosis, i.e., image processing for classification purposes. The main activities are related to the selection and extraction of the right features, in order to recognize the difference among various images. Some pre-processing steps should be considered also, as image scaling and registration.

The main processing blocks are mainly for signal processing, i.e., data acquisition and pre-processing, time-frequency transforms, image processing, classification and diagnosis, and are described in the Sections of the paper. Direct classification of time-frequency images does not offer always the best results in CDD activities, as described in [13] [14]. It is the main objective of this paper to define algorithms for feature selection and extraction, in order to obtain better results in the future classification and diagnosis stages.

The rest of this paper is organized as follows. Section II describes the basic transforms applied to the vibrating signals, i.e., time-frequency and Renyi entropy. Section III describes the basic structure of the proposed method, including data description and time-frequency images. Section IV goes into the results of the experiments, where the main results and examples are presented and discussed. The conclusion and acknowledgement close the article.

II. Data Transforms

The signal under transform is generated by a sliding window with a length depending on the dynamic properties of the analyzed signal, as in the Figure 2. Signal transforms are used to compute specific features of the analyzed signal or to change the analysis system, e.g., time-frequency transforms, or to compute and extract other relevant features, e.g., the Renyi entropy. The final goal of these processing steps is to detect changes in the data stream, preferably associated with individual or mixed faults of the elements and components of the process.

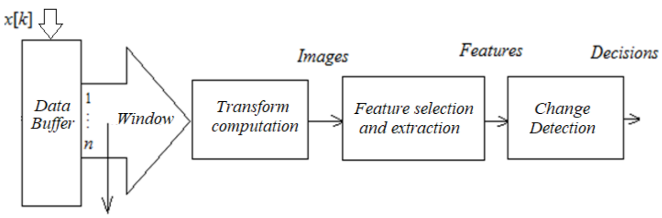


Figure 2. The structure of the processing for feature selection

A. Time-Frequency Transform

Time-frequency transforms are advanced processing techniques for data processing, and especially for data coming from non-stationary signals. A general theoretical framework is presented in [15] [16]. Examples of signals and applications are audio signals [17], mechanical vibrations [18] or biomedical signals [19].

There are three main methods currently used for time-frequency representation and analysis. These are: (i) Short-Time Fourier Transform (STFT); (ii) Wavelet Transform (WT); (iii) Cohen class.

The STFT of a signal $x(t) \in L^2(\mathbf{R})$ considers a window *w(t)*, as

$$STFT_{xw}(t,f) = \int_{-\infty}^{\infty} x(\tau) w(\tau - t) e^{-j2\pi f\tau} d\tau \tag{1}$$

where *w(t)* is the weighting window. The squared modulus is called spectrogram, as

$$S_{xw}(t,f) = \left| STFT_{xw}(t,f) \right|^2 \tag{2}$$

and constitutes a signal energy distribution in the time-frequency plane. Even the spectrogram constitutes one of the widely used methods for the analysis of non-stationary signals, in some case it is unsuitable for the compromise needed of time and frequency resolutions, i.e., it is not possible to simultaneously have good time resolution and good frequency resolution. Consequently, the user must correctly choose the characteristics of the analysis window depending on the signal structure, considering especially the proximity and evolution of the signal components in time and in frequency.

The signal *x(t)* is a function of time, and its STFT is a function of time and frequency. This transform is linear and depends on the chosen window, *w*. Details on how to choose the parameters of the observation window, as length and shape, and the discrete-time version, are presented in [20] [21] [22].

The Wavelet Transform (WT) was promoted to solve the time-frequency resolution problems of Fourier-type methods. A concept called ”multi-resolution” or “multi-scale” is promoted. In the case of continuous time wavelet transform, a basis of translated and dilated functions called wavelets are used as

$$\psi_{t',a}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-t'}{a}\right) \tag{3}$$

The wavelet transform is then

$$WT_{xw}(t,a) = \int_{-\infty}^{\infty} x(\tau) \frac{1}{\sqrt{a}} \psi^*\left(\frac{\tau - t'}{a}\right) d\tau \tag{4}$$

The Cohen class is the set of all bilinear representations, invariant under time and frequency translations, and described by the equation

$$\begin{aligned} C_x(t,f) &= TFR(x(t,f) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} K(t_1,t_2;t,f) x(t_1) x^*(t_2) dt_1 dt_2 \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} k(t,f;v,\tau) x\left(v+\frac{\tau}{2}\right) x^*\left(v-\frac{\tau}{2}\right) dv d\tau \end{aligned} \tag{5}$$

with

$$k(t,f;v,\tau) = K\left(v+\frac{\tau}{2}, v-\frac{\tau}{2}; t, f\right) \tag{6}$$

where the kernel $k(t,f;v,\tau)$ has some special properties, as discussed in [23]. By an equivalent parameterization the equation (5) becomes

$$\begin{aligned} C_x(t,f) & \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi_{t-f}(t-v, f-v) \cdot W_x(u,v) dv d\tau \end{aligned} \tag{7}$$

with

$$\phi_{t-f}(t,\tau) = K(t,0;0,\tau) \tag{8}$$

and

$$W_x(t,f) = \int_{-\infty}^{\infty} x\left(t+\frac{\tau}{2}\right)x^*\left(t-\frac{\tau}{2}\right)e^{-j2\pi f\tau}d\tau \qquad (9)$$

The function $W_x(t,f)$ is called the Wigner-Ville distribution (WVD), [20] [24], being one of the most important members of the Cohen's class. It may be the only distribution with real values that satisfies the properties necessary for the classical applications of signal processing. It is also the only distribution to provide perfect localization for impulse signals and signals with a linearly modulated frequency, [24]. In particular, the WVD is always real-valued; it preserves time and frequency shifts and satisfies the marginal properties. It has also some drawbacks, as the apparition of the cross-terms. This is the reason for using the Choi-Williams Distribution (CWD) [24], where the kernel function is

$$\phi(t,\tau) = \exp\left[-(t\cdot\tau)^2/\sigma^2\right] \qquad (10)$$

This distribution function adopts exponential kernel to suppress the cross-term that results from the components that differ in both time and frequency centers.

The discrete time WVD is defined by [24]

$$\begin{aligned} W(n,m) &= \frac{1}{2N}\sum_{k=0}^{N-1} x(kT) \\ &\cdot x^*((n-k)T)\cdot\exp\left(-\frac{j\pi\cdot m\cdot(2k-n)}{N}\right) \end{aligned} \qquad (11)$$

It is informal to verify that $W(n,m)$ is a periodic function of period $2N$ in both time and frequency. The last relationship shows that in the range $0 \le n < 2N-1, 0 \le m < 2N-1$, representing one period, the WVD needs only be calculated over the range $0 \le n < N-1, 0 \le m < N-1$, having an area of one quarter that of the complete period.

The coefficients of the time-frequency transform define an image, which will be called a Time-Frequency Image (TFI).

B. Entropy Transform

The Renyi entropies are important measures of the information, in wide sense. The measures are scale-dependent when applied to continuous distributions, so their absolute values are meaningless. Therefore, they can generally only be used in comparative or differentiable processes. The information content and the complexity of a probability density function can be measured by this entropy function. The Renyi entropy is intensively used in the field of statistical signal processing, especially in non-stationary conditions, being able to estimate the number of the components of complex signals and the degree of randomness in various signal representation framework, in time or frequency domains [25] [26] [27].

In the case of continuous signal $X(t)$, the Renyi's entropy of the order α is defined as:

$$H_\alpha(X) = -\frac{1}{1-\alpha}\log_2 \int f_X(x)dx, \alpha > 0, \alpha \neq 1 \qquad (12)$$

where $f_X(x)$ is the probability density function (pdf). For univariate discrete signals the common expression for the α-order Renyi entropy is

$$H_\alpha(P) = -\frac{1}{1-\alpha}\log_2 \sum_{j=1}^{N} P_j^\alpha, \alpha > 0, \alpha \neq 1 \qquad (13)$$

As relations (12) and (13) show, the computation of the entropies needs the availability of the exact or estimated pdf. (The probabilities' set, in discrete case). There are estimation solutions based on, e.g., Gaussian kernels, which provides expression as [28]

$$\begin{gathered} \hat{H}_\alpha(X,\sigma) = \frac{1}{1-\alpha}\log \\ \frac{1}{N}\sum_{n=1}^{N}\left(\frac{1}{N}\sum_{k=1}^{N}\left(G(X(n)-X(k), 2\sigma^2\right)\right)^{\alpha-1}, \\ \alpha > 0, \alpha \neq 1 \end{gathered} \qquad (14)$$

The entropy estimators require the selection of the kernel size, σ . This should be small (relative to the standard deviation of the data). Values between 0.1 and 2 for unit-variance signals are good choices, [29].

In the case of images, **I**, a normalized image is considered as a probability density function. The α-order Renyi entropy from [25] is considered as

$$HR_\alpha(\mathbf{I}) = -\frac{1}{1-\alpha}\log_2 \iint\left(\frac{I(t,f)}{\iint I(u,v)dudv}\right)^\alpha dtdf \qquad (15)$$

By discretization of this measure, i.e., by setting $t = n\cdot\Delta t$ and $f = k\cdot\Delta f, n,k \in Z$, it results an expression as

$$\begin{aligned} HR_\alpha(\mathbf{I}) &= -\frac{1}{1-\alpha}\log_2 \sum\sum\left(\frac{I[n,k]}{\sum\sum I[n',k']}\right)^\alpha \\ &+ \log_2(\Delta t\cdot\Delta f) \end{aligned} \qquad (16)$$

By comparing the entropy of two images, associated to consecutive frames – as evolution in time, is possible to detect the differences and thus to make change detection.

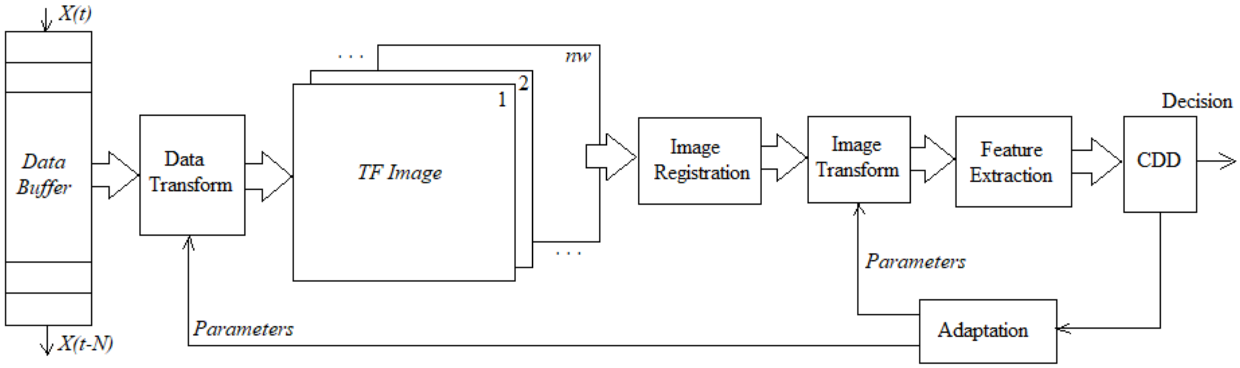


Figure 3. The general structure of the method for selection and extraction of features

III. DESCRIPTION OF THE METHOD

Considering some results obtained by other previous studies and works, e.g., [13], a method to select and to extract the features of the time-frequency image and to define a new set of features, is developed.

The structure of the method is presented in Figure 3. Data coming from environment/process is stored in a data buffer for analysis and processing. Depending on objective, a data transform, e.g., Choi-Williams time-frequency transform, is applied to obtain an image. Depending on processing resources and time-constants of the analyzed signals and process, a set of TFIs is obtained. The set of these images is then registered based on the detection and computation of the main components in the analyzed image. The registered image is the base for feature selection process. The selected features could have physical meaning, as frequency, energy, bandwidth or spreading, or might be generated from other image transforms, as, e.g., Hough transform [30] [31] or Discrete Cosine Transform [32] [33]. In this work, a transform based on the representations of isolines (line of equal elevation) of a matrix is applied [34][35]. The features of these contours, as number, area, average, etc. could be used as feature for CDD objectives. Both data transforms, i.e., for time-frequency image and feature generation need parameters. The number and the values of these parameters depend on CDD performance/results. Thus, an adaptation block is necessary in order to establish the necessary parameters for each transform, as presented in Figure 3.

A. Vibration Data

Data were considered for the case of faults in bearings, available from [36], which are also well explained and analyzed in [37]. The number inside of the round parenthesis indicates the names of the files from the original source of data vibrations, i.e., [36]. Data are briefly described in Table I. Three types of faults are available, like F1 (Inner race), F2 (Ball) and F3 (Outer race). The case F0 means no faults. In the case of the fault F3, there are three sub-cases, depending on the fault position relative to the load zone: ‘centered’ (fault in the 6.00 o’clock position), ‘orthogonal’ (3.00 o’clock) and ‘opposite’ (12.00 o’clock), [37].

TABLE I. DATA TEST SET

Fault size	Faults in bearings			
	F1	***F2***	***F3***	***F4***
	Free	***Inner Race***	***Ball***	***Outer Race***
0.000 “	**d0** (97) (case#0)	-	-	-
0.007"	-	**d1**(105) (case#1)	**d2** (118) (case#2)	**d3** (130) (case#3)
0.014"	-	**d6** (169) (case#6)	**d7** (185) (case#7)	**d8** (197) (case#8)
0.021"	-	**d9** (209) (case#9)	**d10** (222) (case#10)	**d11**(234) (case#11)
0.028"	-	**d14** (3001) (case#14)	**d15** (3005) (case#15)	-

Vibration data from four sizes of the faults are available, data having the advantage of consistency, by considering faults from incipient/small size (0.007") to larger (0.028").

The sampling rate is 12,000 Hz, the motor is with no load, and all data are from drive end bearing (DE). A set of four classes of patterns are considered as: C#0 – no faults, defined by d0; C#1 –inner race faults, defined by {**d1, d6, d9, d14**}; C#2 – ball faults, defined by {**d2, d7, d10, d15**}; C#3 - outer race faults, defined by {**d3, d8, d11**}. The vectors **d4, d5, d12**, and **d13** correspond to other sites of the transducers. The set of classes are defined as

Class #1: **d0** // *Fault free* (17)

Class #2: **d1, d6, d9, d14** // *Fault* 1 (18)

Class #3: **d2, d7, d10, d15** // *Fault* 2 (19)

Class #4: **d3, d8, d11** // *Fault* 3 (20)

For tests based on computer simulation, some new names for a variable were considered. All names beginning with "**d**" indicates a vector with 5,000 samples from normal conditions (no faults) and 5,000 samples from the records with faults. The variable **d0** contains the first 5,000 elements of the raw file named 97 from [36].

Figure 4 presents a sample of time varying signals from each considered class, by considering the first 1,000 samples from each file. The signals are scaled to [-1,1] by normalization. The waveforms seem to be quite different.

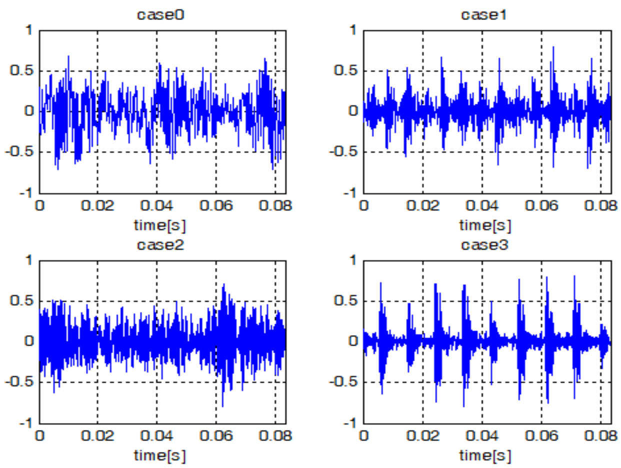


Figure 4. The structure for feature selection

B. *Time-Frequency Images (TFI)*

Figures 5-8 present a set of TFIs, one for each class. Frame or window no 5 is considered for all data records. There are also presented the time evolution, on the bottom side, from 0 to 0.08 [s], and the power spectral density, on the left side of each figure. For the case with no faults, i.e., Figure 5, the spectrum is centered roughly on 1,000 Hz. For the other cases, which cases with faults, the power spectrum density (psd) is spreading up to 4,000 Hz. The shape of the psds indicates some periodic components, like in the Figures 5-8, but also shapes close to the spectrum of modulated signals. These are signals with high frequency bandwidth, with spectral components from 500 Hz up to 4,000 Hz, like in Figure 6. These cases could generate real difficulties in processing and – later - in the detection and classification blocks.

In order to compute a prototype image for each class, the TFIs associated to data frames should be registered, and finally averaged to obtain a prototype image for each class. This is based on a stationarity hypothesis of the processed signals, and thus the main components from TFIs can be registered in both directions, horizontal and vertical. In order to keep the physical meaning/positions of these components, the registration techniques are restricted to horizontal translations only. The next subsection presents some aspects and results from the registration stage.

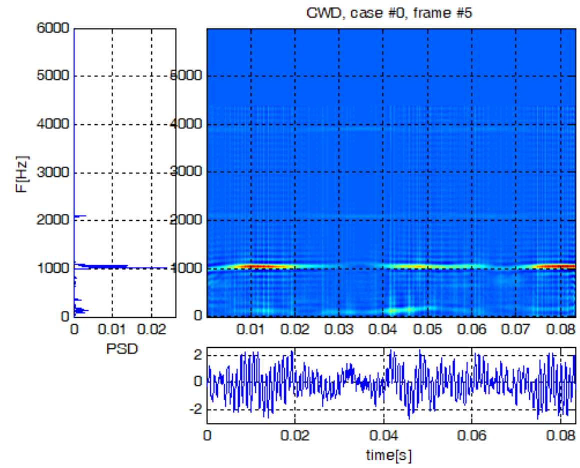


Figure 5. Time frequency image, Class #1 (free of faults)

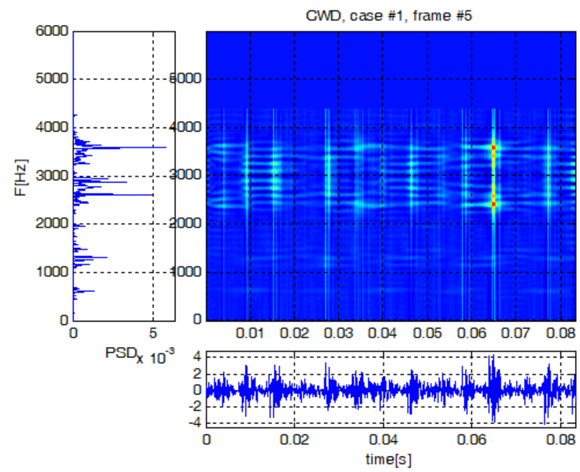


Figure 6. Time frequency image, Class #2

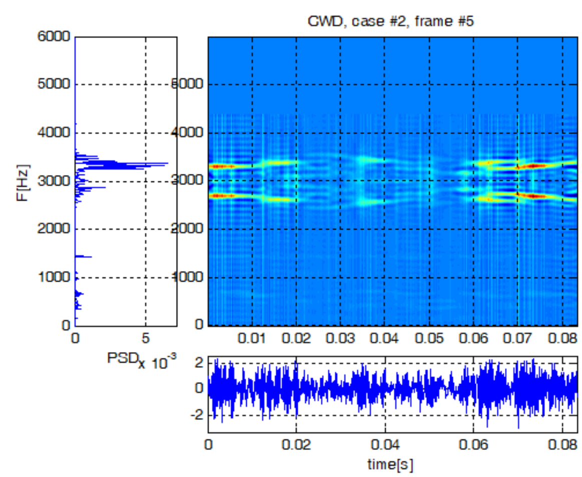


Figure 7. Time frequency image, Class #3

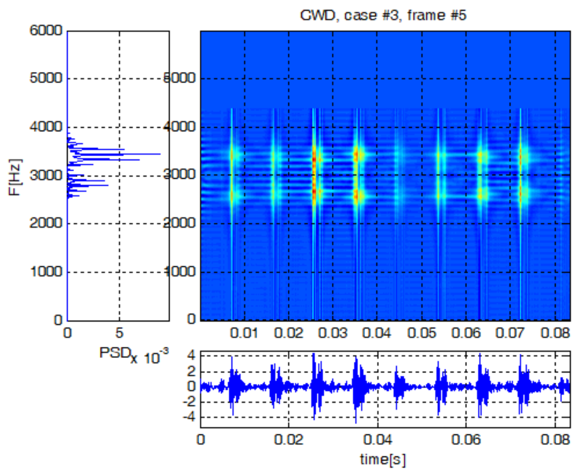


Figure 8. Time frequency image, Class #4

C. Image registration

The registration process of the images is a pre-processing step, before the computation and the classification of the prototype images.

Common algorithms for image registration could use translation, on x and y directions; rigid processing, which means translation plus rotation; similarity, which means translation, rotation and scaling; affine transformation, which considers translation, rotation, scaling, and shearing. The choice of one of them is based mainly on the content of the image, the sources and the number of the images, which are considered for registration. Simple registration methods of the images, from the content point of view, use intensity-based registration algorithms. As complexity rises, the feature-based method is more indicated. Details and examples are available in many references as in, e.g., [38] [39].

The registration time is rapidly growing from translation to affine transformation. Sometimes, for complex transforms - like affine, the registration process could diverge. This is the reason to consider new methods valid for time-frequency images – in general – and in the case of bearings. The registration of TFI has only one degree of freedom, in the sense that any TFI processing must preserve the information of vertical axis, i.e., the frequency axis.

An adapted procedure considers several maxima from time-frequency image, which are considered as references, i.e., their positions remain unchanged during and after registration. Thus, the registration considers the physical meaning of these components, which should have the physical parameters (e.g., frequency), whatever the moment on the time axis. There should be also a distortion limit, in the sense that all images, which are far of the reference, should be removed from the registration set. More details are available in [40]. Further works might investigate some filtering techniques, in order to eliminate the noise or other components.

The reference and the registered images, in pair for each class, are presented in the Figure 9, for a window length of 2,000 samples.

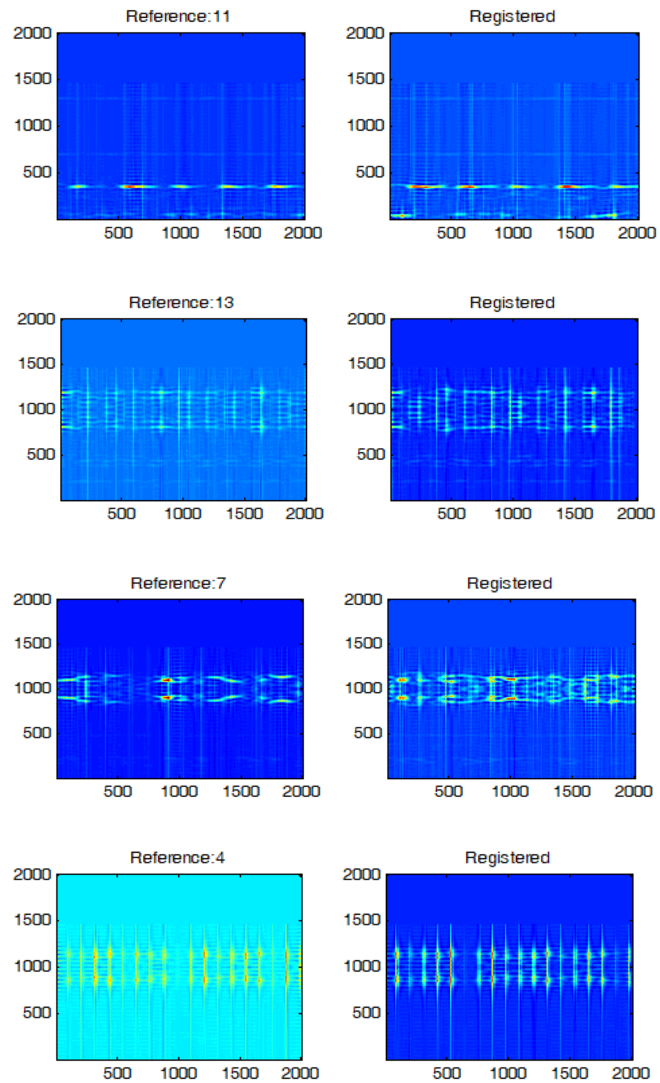


Figure 9. The referenece and the registered images, for classes #1 to #4

D. Feature selection

All images are quite complex, by having many components with various shapes. A first exploratory idea in order to describe the complexity of these images is to compute the Renyi entropies. The numeric results are presented below by the Table II with bold numbers and the averages values of the classes are represented in Figure 10, which does not show important differences among them – especially for the case of faults. As example, cases faults F2 and F4 are difficult to distinguish, the difference between entropies being 0.02 only. The range values of the Renyi entropies indicate the possibility of change detection but difficulties for classification.

An improvement in describing more accurate the content of the TFIs is to consider the shape of the main peaks from the analyzed image. A pre-defined number of peaks could be considered, e.g., 1 to 3, depending on the complexity of the image. Thus, a new image is considered and defined in terms of contours, defined by the above peaks, which will be called transformed image or contour-based image (CBI). This transformation reduces the computation task, by keeping the information about the shape and position of the main components.

TABLE II. THE RENYI ENTROPIES OF THE REGISTERED IMAGES

	Faults / Classes			
	F1	***F2***	***F3***	***F4***
	Free	***Inner Race***	***Ball***	***Outer Race***
0.000 "	**4.40** (case#0)	-	-	-
0.007"	-	**6.08** (case#1)	**5.93** (case#2)	**5.64** (case#3)
0.014"	-	**4.25** (case#6)	**5.89** (case#7)	**6.18** (case#8)
0.021"	-	**4.60** (case#9)	**5.77** (case#10)	**4.09** (case#11)
0.028"	-	**6.24** (case#14)	**5.30** (case#15)	-
Average	**4.40**	**5.29**	**5.72**	**5.31**

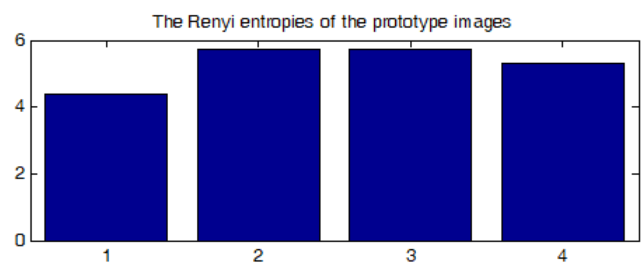


Figure 10. The averaged Renyi entropies of the registered images

In the set of the next two figures, i.e., Figure 11 and Figure 12, the raw/original images and the transformed images are presented, for classes *C*#1 and C#2. On blue background, the set of the registered images are presented. The registered images are considered the prototypes of the classes, from pattern recognition point of view.

The transformed images are presented on white background. Some details are presented in Figure 13, for three values of the number of contours (*nc* =1, 2, and 3). As the number of contours is rising, the shape is coming more complex. A primary analysis of the transformed images reveals some interesting properties:

a) the common content of the images is of vertical curves, as C#1(1,6,9), C#3(3,11,12);
b) the class C#2 has a very complex pattern, for all cases (2,7,10, and 15);
c) the classes C#1 and C#3 have some strange patterns, C#1(14) and C#3(8). Keeping all these images will damage the final classification.
d) the vertical curves/contours of the transformed images are in fact a set of contours, which could be described each by numbers and areas. These could be used as features of the CBIs and later for the associated fault.

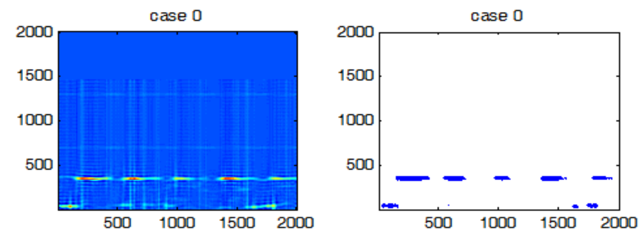


Figure 11. Original and transformed image, Class #1

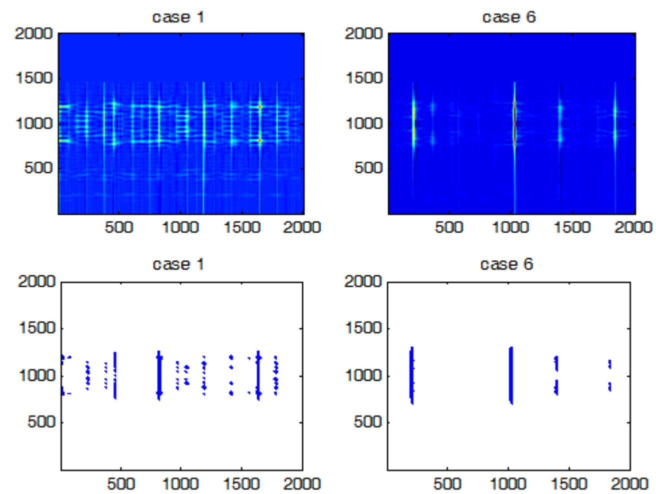


Figure 12. Original and transformed images, class #2 and #3.

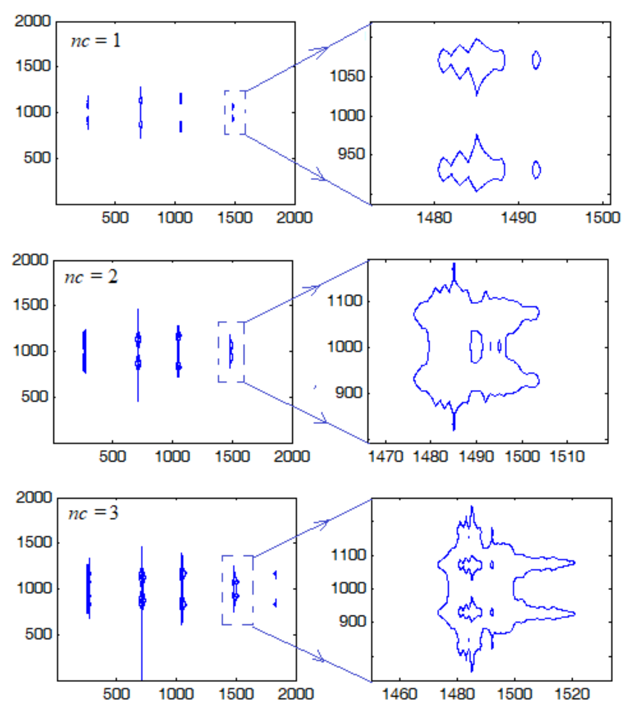


Figure 13. Details of contours based images, *nc* = 1,2. 3.

In order to extract the right information, as features from the transformed images, some elements or parameters should be considered and properly used. In this work, parameters based on the number and size of the contours are used, mixed with the values of some statistical moments.

The basic elements of the feature vector are:

i) the number of contours, *Nc*, as a measure of the complexity;
ii) the area of the polygons, *Ac*, as a measure of the spreading on horizontal plane;
iii) the variance of the above areas, *var(Ac)*, as a measure of the complexity;
iv) the average of the area of the polygons, $E\{A_c\}$=Ac;
v) the mean of the squared values of areas, $E\{A^2{}_c\}$;
vi) the Renyi entropy of transformed images, *RH*.

A vector of features is defined for each class by using the above features, as

$$\mathbf{f}_i = \left[N_c \quad \sum A_c \quad \mathrm{var}(Ac) \quad \overline{A_c} \quad \overline{A_c^2} \quad RH \right], \quad i = 1,2,3,4 \tag{21}$$

For each data vector, **d**, from a class, the vector $\mathbf{f}_i$ is evaluated, and matrix of features is obtained for each class, as

$$\mathbf{F}_j = [\mathbf{f}_1 \quad \mathbf{f}_2 \quad \ldots \quad \mathbf{f}_4],\ j = \overline{1,4} \tag{22}$$

The effect of the features is estimated by a general discriminant matrix of the classes

$$D(k, j) = \sum_{k=1}^{4} \sum_{j=1}^{4} (\mathbf{F}_k - \mathbf{F}_j)^2, k, j = \overline{1,4} \tag{23}$$

or, by considering only the distinct classes, by the discriminant function

$$D_1 = E\{D(k, j) \mid k \neq j, k > j, k, j = \overline{1,4}\} \tag{24}$$

or

$$D_2 = \sum_{k=1}^{4} \sum_{j=1, k \neq j}^{k} D(k, j) \tag{25}$$

This should be high as possible. The number of the features considered in the Equation (21) could be modified, in order to gain the highest dissimilarity among classes.

IV. Results of the Experiments

The evaluation of the features for all classes is presented in Figure 14, with green for C#1, yellow for C#2, for C#3 and black for C#4. A high variance of the features for patterns of the same class is observed, e.g., feature 1 for class #2, #3, and #4; feature 2 for classes #3 and #4; feature 6 for classes #2 and #3. If the variance is associated with the size of the fault, then a criterion to select the right features is to maximize the dissimilarity among classes.

The mean values of the features are presented in Figure 15, with different colors for classes. There six features for each class of four. There is a difficulty to make a good classification, especially for the classes 2, 3 and 4, where the evolution and the range of the values seem quite close. A solution to change this is to increase the number of contours (*nc*) in CBIs from 1 to 2, and 3, with the results presented in Figure 16. Based on these evolutions, the next step is to evaluate a criterion for the selection of the best features for classification purposes, i.e., to maximize the dissimilarity of the classes in the feature space. Figure 17 shows the evolution of the dissimilarities in terms of the six features and based on Equation (23).

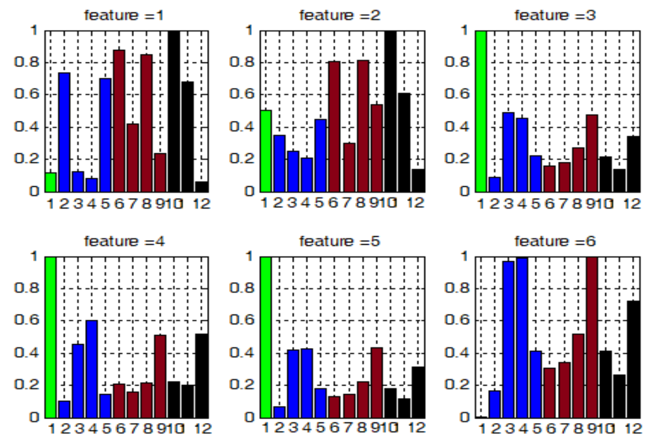


Figure 14. Feature space among various cases, from 1 to 12

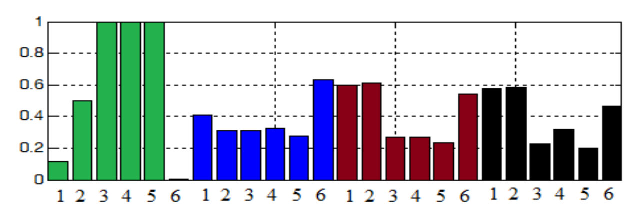


Figure 15. The mean values of the features; *nc* =1

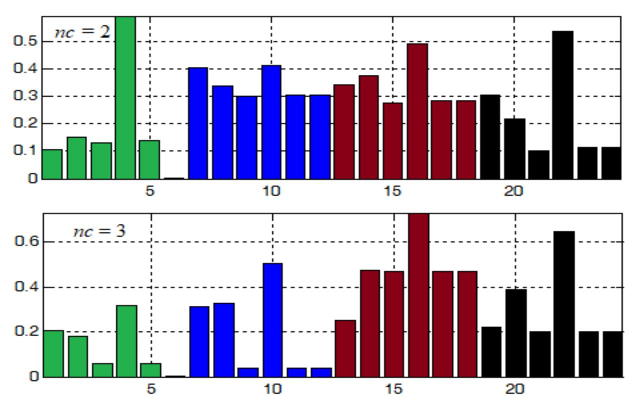


Figure 16. The mean values of the features; *nc* =2, and 3

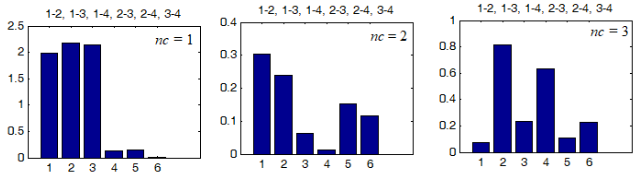


Figure 17. Dissimilarities among classes

The values of the discriminant functions (24) and (25) for various values of the number of contours are presented in Table III. The highest value of the discriminant functions is obtained for one contour only, *nc* = 1.

TABLE III. VALUES OF THE DISCRIMINANT FUNCTIONS

	nc				
	1	2	3	4	5
D_1	1.094	0.147	0.349	0.375	0.481
D_2	6.596	0.884	2.094	2.251	2.888

CONCLUSION

The objective of the paper was to promote a method for feature selection and extraction from time-frequency images, as an alternative to some classical well-known methods, as those based on Hough Transform or Discrete Cosine Transforms.

Experiments used real vibration data coming from bearings of the rotating machines, bearings with various faults and sizes. The proposed method is general and can be applied also to other types of data, mechanically generated or not.

The roots of the method come from the fact that for classification purpose, the complexity of time-frequency images is not properly described by Renyi entropy. More information in terms of more features must be considered at the input of the classifier.

The method uses two data transforms. The first one is based on Choi-Williams time-frequency transform and the second uses a representation based on isolines of a matrix, applied to the main components of the time-frequency images. Before extracting the features, the time-frequency images are registered. Depending on the number of the contours obtained, which could vary, e.g., from 1 to 5, the features are varying and change the dissimilarity of the classes.

An important step is adaptation of the parameters for the used data transform. Depending on the dynamic properties of the process which generates the mechanical vibrations and depending on the evolution of the faults, the user must check the length of the data frame/window, the number of frames, the numbers of the contours which define the transformed images (contour – based).

The selected features are based on the statistical moments, as average, variance and squared average values of the areas of the contours. The information- based features is also used, by considering the Renyi entropy. Larger feature vectors could be considered by including also the number of the main components and their centers, in time and space.

The results show a good separability in the feature space, in the sense of clustering, and thus the possibility to obtain efficient classifiers.

The method could be extended to more complex signals and applications. The feature vector could be extended with qualitative or quantitative parameters, which describe the shape of contours, i.e., the distance to some standards cures as circles, ellipses or squares. Further research could consider also fuzzy logic in the description and selection of the features from time-frequency images.

ACKNOWLEDGMENT

The work was partly supported by the Romanian Council for Research (UEFISCDI) under Grant 224/2014, "Experimental model for change detection and diagnosis of vibrational processes using advanced measuring and analysis techniques model-based (VIBROCHANGE)".

REFERENCES

[1] D. Aiordachioaie, Th. D. Popescu, and B. Dumitrascu, "A Method of Feature Extraction from Time-Frequency Images of Vibration Signals in Faulty Bearings for Classification Purposes", Proc. of International Conference on Emerging Networks and Systems Intelligence (EMERGING '2019), 22-26 September, 2019, Porto, Portugal.

[2] J.H. Williams, A. Davies, and P. R. Drake (Eds), Condition-based Maintenance and Machine Diagnostics, Springer, 1994.

[3] L. Fedele, Methodologies and Techniques for Advanced Maintenance, Springer, 2011.

[4] M. Bengtsson, E. Olsson, P. Funk, and M. Jackson, "Technical Design of Condition Based Maintenance System", Proceedings of the 8th Congress, University of Tennessee – Maintenance and Reliability Center, Knoxville, USA, May 2nd – 5th, Paper III, 2004.

[5] K. Shin and J. K. Hammond, Fundamental of Signal Processing for Sound and Vibration Eng., John Wiley & Sons Ltd, 2008.

[6] D. M. Yang, A. F. Stronach, and P. MacConnell, "The Application of Advanced Signal Processing Techniques to Induction Motor Bearing Condition Diagnosis", P. Meccanica, Springer, 2003, vol 38(2), pp 297-308.

[7] D. Mironovs and A. Mironov, "Vibration based signal processing algorithm for modal characteristics change assessment", AIP Conference Proceedings 2029, 020043, 2018, pp.1-10.

[8] B. Boashash (Ed), Time-Frequency Signal Analysis and Processing, Elsevier, 2016.

[9] L. Stankovic, M. Dakovic, and T. Thayaparan, Time-Frequency Signal Analysis with Applications, Artech House, 2014.

[10] M. Boufenar, S. Rechak, and M. Rezig, "Time-Frequency Analysis Techniques Review and Their Application on Roller Bearings Prognostics". In: Fakhfakh T., Bartelmus W., Chaari F., Zimroz R., Haddar M. (Eds) Condition Monitoring of Machinery in Non-Stationary Operations. Springer, Berlin, Heidelberg, 2012.

[11] B. Boashash, L. Boubchir, and G. Azemi, "A methodology for time-frequency image processing applied to the classification of non-stationary multichannel signals using instantaneous frequency descriptors with application to newborn EEG signals", EURASIP Journal on Advances in Signal Processing, Article 117, 2012.

[12] D. Rossetti, Y. Zhang, S. Squartini, and S. Collura, "Classification of bearing faults through time-frequency

analysis and image processing," 2016 17th International Conference on Mechatronics - Mechatronika (ME), Prague, 2016, pp. 1-7.

[13] D. Aiordachioaie, N. Nistor, and M. Andrei, "Change Detection in Time-Frequency Images by Feature Processing in Compressed Spaces", IEEE 24th Int. Symp. for Design and Technology in Electronic Packaging (SIITME), Iași, Romania, 2018, pp.181-186 .

[14] E. Sejdic, I. Djurovic, and J. Jiang, "Time–frequency feature representation using energy concentration: An overview of recent advances", Digital Signal Processing, 19, 2009, pp. 153-183.

[15] Q. Zhu, Y. Wang, and G. Shen, "Comparison and Application of Time-Frequency Analysis Methods for Nonstationary Signal Processing". In: Lin S., Huang X. (eds) Advanced Research on Computer Education, Simulation and Modeling. CESM 2011. Communications in Computer and Information Science, vol 175. Springer, Berlin, Heidelberg, 2011.

[16] F. Hlawatsch and F. Auger, "Time-Frequency Analysis. Concepts and methods", ISTE Ltd and John Wiley & Sons, 2008.

[17] M. Karjalainen and V. Pulkki, Communication Acoustics: An Introduction to Speech, Audio and Psychoacoustics, John Wiley & Sons, 2015.

[18] F. Al-Badour, M. Sunar, and L. Cheded, "Vibration analysis of rotating machinery using time–frequency analysis and wavelet techniques", Mechanical Systems and Signal Processing, vol. 25 (6), 2011, pp. 2083-2101.

[19] M. Akay, Time Frequency and Wavelets in Biomedical Signal Processing, Wiley-IEEE Press, 1997.

[20] B. Boashash, "Introduction to Time-Frequency Signal Analysis". 10.1007/978-1-4612-0137-3_11, 1997.

[21] L. Cohen, "Time-Frequency Analysis", Prentice Hall, 1995.

[22] J. Ville, "Theorie et applications de la notion de signal analytique", Câbles et transmissions, vol. 2A,1948, pp. 66‑74.

[23] H. Choi and J.W. Williams., "Improved time-frequency representation of multicomponent signals using exponential kernels", IEEE Trans. Acoust., Speech, Signal Process., vol. 37, no. 6, 1989, pp. 862–871.

[24] P. D. McFadden and W. Wang, Time-Frequency Domain Analysis of Vibration Signals for Machinery Diagnostics. (I) Introduction to the Wigner-Ville Distribution, University of Oxford, Report OUEL 1859/92, 1990.

[25] P.A. Bromiley, N.A. Thacker, and E. Bouhova, "Shannon Entropy, Renyi Entropy, and Information", Statistics and Information Series (2004-004), 2004.

[26] R. G. Baraniuk, P. Flandrin, A. J. E. M. Janssen, and O. J. J. Michel, "Measuring Time–Frequency Information Content Using the Rényi Entropies", IEEE Transactions on Information Theory, vol. 47(4), 2001, pp. 1391-1409.

[27] D. Aiordachioaie, "Some properties of Entropies for Change Detection", The Annals Of "Dunărea De Jos" University Of Galati Fascicle III, 2013, Vol. 36, No. 1, ISSN 1221-454x Electrotechnics, Electronics, Automatic Control, Informatics, 2013, pp.43-48.

[28] D. Erdogmus K.E.Hild II, and J.C. Principe, "Blind source separation using Renyi's alpha-marginal entropies", Neurocomputing, vol. 49 (1), 2002, pp. 25–38.

[29] K.E. Hild II, D. Erdogmus, J.C. Principe, "An analysis of entropy estimators for blind source separation", Signal Processing, vol 86 (1), 2006, pp.182-194.

[30] V.F. Leavers, Shape Detection in Computer Vision Using Hough Transform, Springer, 1992.

[31] A. M. Tekalp, Hough Transform Methods, in Handbook of Image and Video Processing (Second Edition), Al Bovik Editor, AP, 2005.

[32] K.R. Rao and P. Yip, Discrete Cosine Transform. Algorithms, Advantages, Applications, AP, 1990.

[33] C. Loeffler, A. Ligtenberg, and G. Moschytz, "Practical Fast 1-D DCT Algorithms with 11 Multiplications", Proc. Int'l. Conf. on Acoustics, Speech, and Signal Processing, 1989 (ICASSP `89), pp. 988-991.

[34] D. Hearn and M.P. Baker, Computer Graphics, Prentice-Hall, Inc., 1986.

[35] M. Trott, The Mathematica GuideBook for Graphics, Springer-Verlag New York, 2004.

[36] Case Western Reserve University, Bearing Data Center, At: http://csegroups.case.edu/bearingdatacenter/home , Retrieved: 06, 2019.

[37] W. A. Smith and R. B. Randall, "Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study", Mechanical Systems and Signal Processing, 64-65, pp. 100–131, 2015

[38] G. A. Ardeshir, Image Registration Principles, Tools and Methods, Springer, 2012.

[39] I. N. Bankman (Ed.), Handbook of Medical Image Processing and Analysis, Elsevier, 2009.

[40] D. Aiordachioaie and S. M. Pavel, "Qualitative Analysis of the Time-Frequency Images of Vibrations in Faulty Bearings", IEEE 11th Int. Conf. on Electronics, Computers and Artificial Intelligence (ECAI), Romania, 2019, pp. 1-6.

Communicating Semantic Content to Persons with Deafblindness by Haptograms and Smart Textiles: Theoretical Approach and Methodology

Sándor Darányi, Nasrine Olson[1], Eva Lindell, Nils-Krister Persson
University of Borås
Borås, Sweden
E-mail: {sandor.daranyi, nasrine.olson, eva.lindell, nils-krister.persson}@hb.se

Marina Riga, Efstratios Kontopoulos, Ioannis Kompatsiaris
Information Technologies Institute (ITI/CERTH)
Thessaloniki, Greece
E-mail: {mriga, skontopo, ikom}@iti.gr

***Abstract*— By means of a proof-of-concept prototype, which is work in progress, we adopted a multidisciplinary approach to develop a smart-textile-based communication system for use by people with deafblindness. In this system, sensor technologies and computer vision are used to detect environmental cues such as presence of obstacles, faces, objects, etc. Focusing on the communication module here, a new ontology connects visual analytics with the user to label detected semantic content about objects, persons and situations for navigation and situational awareness. Such labelled content is then translated to a haptogram vocabulary with static vs. dynamic patterns, which are mapped to the body. A haptogram denotes a tactile symbol composed over a touchscreen, its dynamic nature referring to the act of writing or drawing. A vest made of smart textile, in the current variant equipped with a 4 x 4 grid of vibrotactile actuators, is used to transmit haptograms on the user's back. Thereby system messages of different complexity -- both alerts and short sentences -- can be received by the user, who then has the option to respond by pre-coded questions and messages. By means of grids with more actuators, displays with higher resolution can be implemented and tested, paving the way for an extended haptogram vocabulary, covering more detailed ontology content.**

Keywords - deafblind communication; haptograms; word and sentence semantics; ontology; smart textiles.

I. INTRODUCTION

Communication with and between people with deafblindness is constrained by the nature of the condition and the lack of supportive tools and societal structures. This study explores novel communication modes for this group.

Deafblindness refers to a unique combination of vision and hearing loss, where the level of loss in either senses is such that it does not allow compensation of one impaired sense by the other. This condition in its various forms, ranging from congenital (i.e., present at birth or acquired prior to language development) to acquired (i.e., due to illness, accident or ageing), involves a broad diversity in abilities. Based on [1], we report work in progress that aims to address communication challenges that accompany deafblindness. Our approach combines a simple conceptual language, with navigation, situational awareness and communication components. The communication is conducted by means of haptic stimuli, projected on appropriate parts of the human body via a smart textile screen, and is intended to be useful across the complete spectrum of this condition.

Our research is carried out in the SUITCEYES (https://suitceyes.eu/) EU Research and Innovation Action project, with the mission to deploy a prototype which is wearable, combines situational awareness, visual analytics and communication by a joint ontology, and works in distance mode by wireless connection as a default. Instead of a combination of elements of touch, e.g., consecutive dots, dashes and strokes by hand to encode characters and their sequences, we propose to use haptograms, where the limited size and resolution of a body part as screen is counterbalanced by evolving patterns, i.e., the dynamics of signs. Our effort is in line with the approach by [2], building on their Tactile Brush findings. However, the focus here is on language design by means of an ontology-compliant vocabulary vs. grammar, where the latter implements relational contextualization and sign sequencing. Thus, our work belongs to the category of a priori defined spatial-temporal patterns in the semiotic vein. Here we will discuss the integration of ontology, haptograms and textile aspects only from a theoretical and methodological perspective.

The rest of the paper is structured as follows: Section II reviews research in haptics relevant to our design considerations, contrasted with its respective linguistic underpinnings. Section III introduces the SUITCEYES ontology that plays the role of the unified model for semantic integration of information from the environment, while Section IV presents our approach for designing the haptogram vocabulary. Section V discusses implementation in smart textile, whereas Section VI concludes the paper and refers to future work directions.

II. RELATED RESEARCH

A. Research in haptics relevant to our design

For haptogram development we were inspired by [3], and [4], where the authors' work expands over multiple decades. Their approach, Social Haptic Communication (SHC), basically reproduces ideograms on different regions of the body by a combination of hand strokes, gestures, pressure, etc. For this, they developed a rich set of tactile signs, compiled into a simple language with its own syntax built from *haptemes*, i.e., building blocks of touch, and vocabulary of so-called *haptices*, tailored to a range of situations and topics of great practical importance for the receiver. At the same time, due

[1] Corresponding author

to its consensual nature, this approach is idiosyncratic, and in its social mode unable to be applied at a distance.

The above phonemic approach in [4] finds support from [5], where the authors found that decomposing language into phonemes that are transcribed into unique vibrotactile patterns enables people to receive lexical messages on the arm. A potential barrier to adopting this new communication system is the time and effort required to learn the association between phonemes and vibrotactile patterns. However, their study was limited to the learning of 100 patterns by different methodologies, displayed on the arm, and the concepts were not connected to an ontology.

On the other hand, in [6], the authors experimented with a new tactile speech device based on the presentation of phonemic-based tactile codes. The device consisted of 24 tactors, under independent control for stimulation at the elbow to wrist area. Using properties that included frequency and waveform of stimulation, amplitude, spatial location, and movement characteristics, unique tactile codes were designed for 39 consonant and vowel phonemes of the English language. The participants, 10 young adults, were then trained to identify sets of consonants and vowels, before being tested on the full set of 39 tactile codes.

In [7], the authors investigated several haptic interfaces designed to reduce mistakes in Morse code reception of 12 characters. Results concluded that a bimanual setup, discriminating dots/dashes by left/right location, reduced the amount of errors to only 56.6% of the errors compared to a unimanual setup that used temporal discrimination to distinguish dots and dashes.

Very much in line with what we would like to achieve, the authors in [2] proposed Tactile Brush, an algorithm that produces smooth, two-dimensional tactile moving strokes with varying frequency, intensity, velocity and direction of motion. The design of the algorithm was derived from the results of psychophysical investigations of two tactile illusions, *apparent tactile motion* and *phantom sensations*. Combined together they allowed for the design of high-density two-dimensional tactile displays using sparse vibrotactile arrays. In a series of experiments and evaluations, they demonstrated that Tactile Brush is robust and can reliably generate a wide variety of moving tactile sensations for a broad range of applications.

Another related track of research [8] conducted four experiments to evaluate tactile localization and tactile pattern recognition on the torso by means of a one-dimensional eight-tactor display vs. a two-dimensional 16-tactor display to present tactile cues to the waist and back, respectively. They found that a display with eight tactors mounted circumferentially around the waist can provide tactile cues that are perceived very accurately in terms of the location of stimulation, whereas the 16-tactor array on the back was found to be inadequate to support precise spatial mapping, but an array with fewer elements could provide such spatial cues, so that simple navigational and instructional commands can be presented tactually on the torso. These findings were augmented in [9], where the authors evaluated the effectiveness of tactile displays either on the forearm or the back to communicate simple instructions and commands in a military context. Their results suggested that with the judicious selection of tactile patterns both the arm and back provide a functional substrate for tactile communication.

A next source of inspiration was [10], which described a different domain of study. As in a situation of sensory overload, touch is a promising candidate for messaging given that it is our largest sensory organ with impressive spatial and temporal acuity, there is need for a theory that addresses the design of touch-based building blocks for expandable, efficient, rich and robust touch languages that are easy to learn and use; moreover, beyond design, there is a lack of implementation and evaluation theories for such languages. To overcome these limitations, he proposed a unified, theoretical framework, inspired by natural, spoken language, called *Somatic ABC's* for Articulating (designing), Building (developing) and Confirming (evaluating) touch-based languages. To evaluate the usefulness of Somatic ABC's, its design, implementation and evaluation theories were applied to create communication languages for two very unique application areas: audio-described movies, and motor learning. It was found that Somatic ABC's aided the design, development and evaluation of rich somatic languages with distinct and natural communication units.

Implementation-wise, the approach to conveying any tactile code to the body is often rather simple. Glued-on solutions or non-compliant devices are common in many of the more research-oriented studies. For more technical descriptions with users in focus, textiles and garment have been employed and are described by many [11], [12]. The interest in the literature is typically in the hardware and its control system, which soon becomes problematic when the number of actuators and connections increases. Gaming has attracted a lot of interest. By far the most common haptic output device used is the vibrator, denoted vibrotactile element (VTE), with textile aspects being secondary. In [13], the authors provide a solution with pieces made out of neoprene, which is stretchy. Hook-and-loop fasteners were used to keep five large pieces (for left shoulder, right shoulder, left elbow, right elbow, wrist belt) together. Comfort was not discussed, neither was wiring. What was denoted heavy weight Lycra was used in a harness design in [14], in order to achieve tightness and close fit to the body. Assembly and design were not in focus. The authors in [15] used a tight-fitting T-shirt in a study of haptic elements for promoting motion training. Once again, the technical focus was on the hardware. In general, there are few studies where the role of textile and the construction thereof is central.

B. Semantic theories underlying haptograms

Semantic content in language is manifest by means of word vs. sentence meaning. To pin down word semantic vs. sentence semantic theories which can help one identify kinds of meaning implementable for deafblind communication, we took inspiration from Chinese *logograms* representing concepts, instead of characters that stand for speech sounds. They come as a blend from different directions and schools of semantic research [16]. More importantly, one can apply touch primitives to create haptic drawings for concepts, with connections to several semantic theories briefly listed below.

To conceive a set of actuator patterns, which correspond to units in a mental vocabulary, partly overlapping with ontology labels, *semantic primitives* [17] and *semantic universals* [18] were relevant points of departure. These include elementary, archetypical concepts such as substantives (*you, I; someone, people; something*), mental predicates (*think, know, want, feel, see, hear*), descriptors (*big, small*), temporality (*when, after, before, a long time, a short time, now*), etc., while their concatenations can be related to the *Language of Thought hypothesis* [19], a theory which describes the nature of thought as possessing "language-like" or compositional structure, with simple concepts combined in systematic ways (akin to the rules of grammar in language) so that in its most basic form, thought, like language, has syntax. Another aspect to connect semantic primitives as components of concepts with *haptemes* as components of touch [3] originates in structural linguistics [20]; here we chose the interpretation that *phonemes* constitute an abstract underlying representation for segments of words. *Distributional semantics* as in [21], [22], and [23] -- responsible for meaningfulness in dimensional reduction methods and feeding forward to the ontology -- adds another relevant semantic theory to the aforementioned, with statements derived from ontology labels by semantic reasoning going back to *logical semantics* [24]. Another umbrella term for the above in a deafblind context is the research field of *sign language semantics* [25]. A number of relevant PhD theses augmented our resources to relate semantics mapped to the body in a sensory deprivation context, e.g., [26], [27], and [28].

III. THE SUITCEYES ONTOLOGY

The key aim of the SUITCEYES ontology is to semantically integrate information coming from the environment (via sensors), and from the system's analytic components (e.g., visual analysis of camera feed). In this sense, the ontology is primarily focused on semantically representing aspects relevant to the users' context, in order to provide them with enhanced situational awareness, and augment their navigation and communication capabilities. More importantly, the proposed ontology also serves as the bridge between environmental cues and content communicated to the user via the haptograms described in the next section.

In ontology engineering, it is common practice to reuse existing third-party models and vocabularies during the development of a custom ontology, in order to rely on previously used and validated ontologies. We thus adopted the semantic representation of objects and activities from the Dem@Care ontology [29], [30], which contains a set of descriptions of everyday activities and common objects used in an everyday context that are highly relevant to our goals (e.g., mug, plate, toothbrush, furniture, window, door, etc.), all inspired by real case scenarios. Moreover, we relied on SOSA/SSN [31] for representing sensors and the respective observations, and on the Friend-Of-A-Friend (FOAF) specification [32] for representing persons and social associations. Finally, we integrated the SEAS (Smart Energy Aware Systems) Building Ontology [33], which is a schema for describing the core topological concepts of a building, such as buildings, building spaces and rooms. A summary of all imported ontologies is presented in Table 1. The hierarchy of concepts related to rooms and space, adopted from both Dem@Care and the SEAS Building Ontology, is visualized in Fig. 1. The latter set of concepts is grouped under the notion of *sot:SemanticSpace*, where *sot* is the namespace prefix for the SUITCEYES ontology. Moreover, we extended our ontology with additional concepts that emerged during the requirements analysis (see deliverables D2.1 and D2.2 posted on the project website) of our *Haptic, Intelligent, Personalised, Interface* (HIPI), which is a key SUITCEYES outcome.

The ontology, populated with the incoming information (i.e., detected persons, objects etc.), constitutes the system's Knowledge Base (KB). An additional component, the Knowledge Base Service (KBS) that sits "on top" of the KB, ensures that the incoming data are semantically annotated and fused properly into the KB, providing the groundwork for producing a higher-level interpretation of the combined information. The end goal is to enable the HIPI to deliver semantic content to the user with respect to his/her physical surroundings.

TABLE 1. A LIST OF THIRD-PARTY ONTOLOGIES UTILIZED AND EXTENDED WITHIN THE SUITCEYES ONTOLOGY

prefix	**Ontology**	**URL**	**Concepts**	**Imported in SUITCEYES version**
foaf	Friend-Of-A-Friend	http://xmlns.com/foaf/spec/	Person and its asserted properties	v1
sosa	Semantic Sensor Network	https://www.w3.org/TR/vocab-ssn/	Sensor and its asserted properties	v1
dem	Dem@Care	http://www.demcare.eu/results/ontologies	Activity Object Room and their asserted properties	v1 v1 v2
seas	Smart Energy Aware Systems Building Ontology	https://ci.mines-stetienne.fr/seas/BuildingOntology-1.0	BuildingSpace Room and their asserted properties	v2 v2

Finally, a rule-based semantic reasoning mechanism serves specific queries for inferring the context-related information, either in the form of structured content in JSON format, or as natural language phrases.

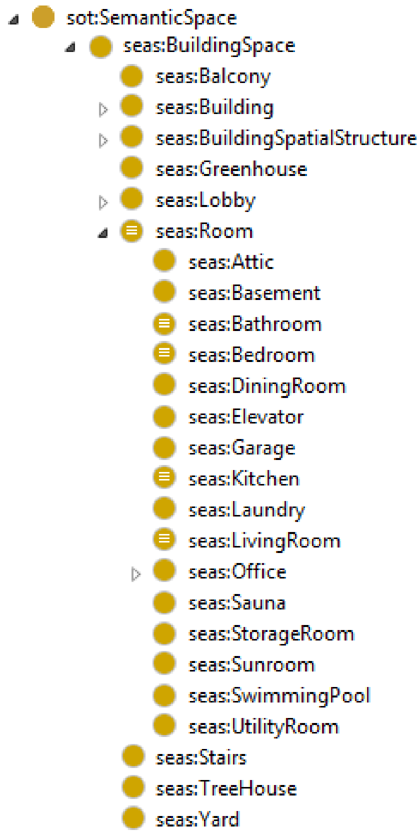


Figure 1. Concept hierarchy related to semantic space, which were adopted in the SUITCEYES ontology from Dem@Care and SEAS Building.

A. *Ontology conceptualization*

Fig. 2 displays an overview of the core ontology classes based on the Graffoo ontology visualization notation [34]: the yellow rectangles represent classes, while the green ones represent data properties (i.e., properties that take a raw data value, like, e.g., integers and strings). The prefixes in front of some of the class names indicate the namespace of the respective third-party ontologies, as mentioned above. Classes and properties that have no prefix belong to the core SUITCEYES ontology.

As indicated in Fig. 2, class *Detection* is fundamental within the context of the SUITCEYES ontology and refers to environmental cues detected by the sensors that have been instantiated in the KB. An instance of type *Detection* may be associated with the relevant sensor(s) that provide data to the ontology, via property *providedBy*. Currently, there are two specific categorizations of class *Sensor* (i.e., *Camera* and *iBeacon*), which are related to the relevant operational sensors attached to the HIPI and provide data to the KB. On the basis of the sensors' data, an instance of class *Detection* can be associated with one or more instances of type *Person* (known or unknown persons), *Object*, *Activity*, and/or *SemanticSpace* (i.e., rooms, building spaces, as previously described); the relevant assertion is achieved via property *detects*. On the basis of the semantic reasoning mechanism mentioned previously, higher-level results that combine incoming data from all sensors, are produced and represented in the ontology via class Output and its specializations: *Alert*, *Message*, and *Warning*.

With regard to spatial relations, we focus on orientation (left/right), existence (in a room) and distance (far/close/immediate). Thus, an entity that occupies space (e.g., persons, objects) is considered a *SpatialEntity*, and the occupied space (e.g., a room or a location) belongs to the *SemanticSpace*. These two aspects formulate the

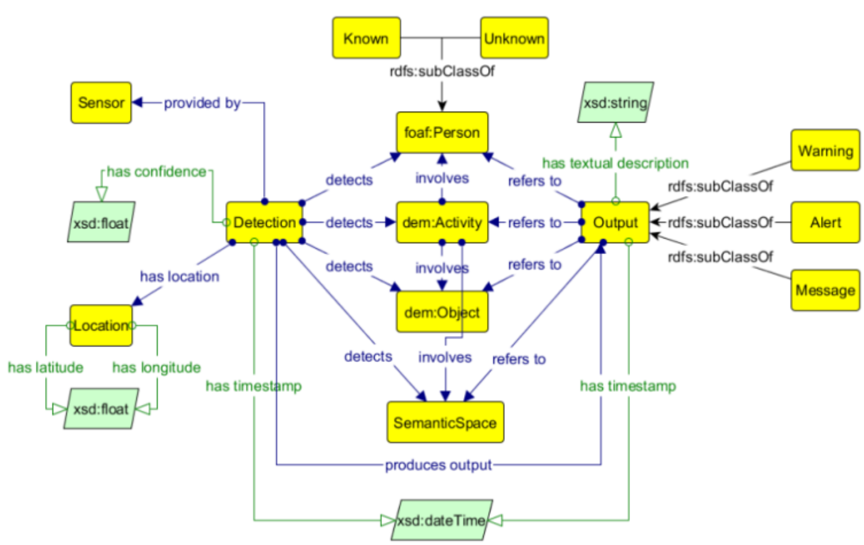


Figure 2. Overview of the core classes of the SUITCEYES ontology.

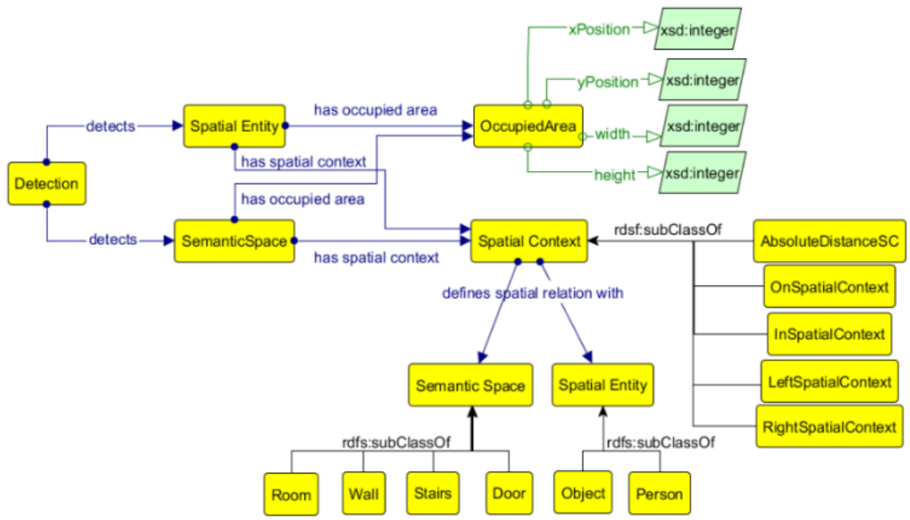


Figure 3. Semantic spaces and spatial contexts in the SUITCEYES ontology.

respective entity's *SpatialContext*, which provides information regarding the entity's relationship to the semantic space it is located in. Examples include *in*, *on*, *left*, *right*, *far*, *close*, etc. The aforementioned concepts are depicted in Fig. 3. These definitions play a key role in the semantic reasoning mechanism, as they form the basis for inferring spatially related information to the user, e.g., which objects are close to or far from the user, what sort of entities are located in the room where the user is, etc.

B. *Sample instantiation*

Based on the ontological concepts presented above, Fig. 4 illustrates a sample instantiation, resembling an activity detected by the system's camera. The activity involves two people speaking to each other: one of them is known to the user (i.e., John) whereas the other is unknown. Moreover, these two people are currently located in the kitchen (i.e., in_room_spatial_context), and the respective message is communicated to the user via a textual description, which is then converted to haptograms as described next.

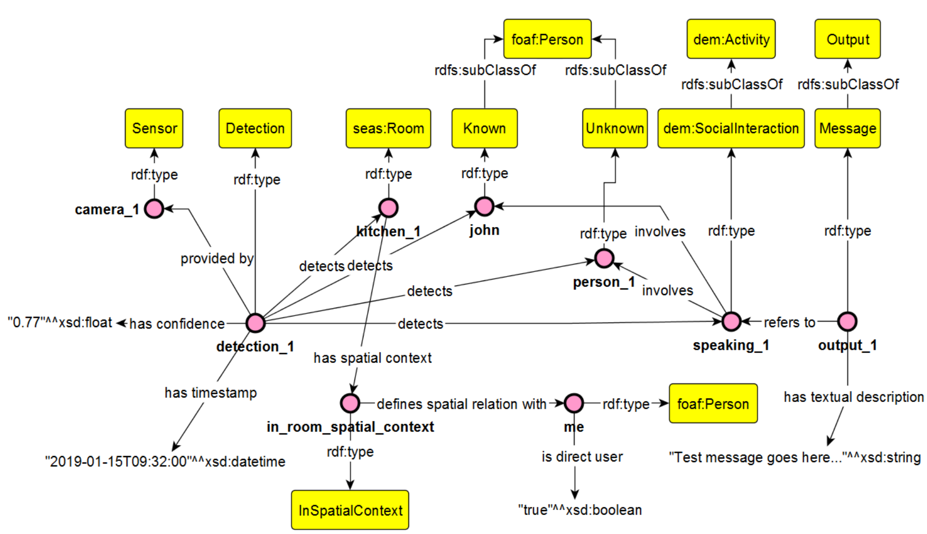


Figure 4. Sample instantiation of an activity involving two people discussing in the kitchen.

TABLE 2. EXAMPLE QUERIES FOR WHICH THE ONTOLOGY CAN PRODUCE RESULTS IN NATURAL LANGUAGE TEXT

#	Query	Answer
Q1	Where is my <phone> now?	Your <phone> is located on your <left>, <close to> you.
Q2	In which room am I now located?	You are located in the <kitchen>.
Q3	When and where was my <mug> detected for the last time?	Your <mug> was found in the <living room>, <10 seconds> ago.
Q4	Which <entity> is observed on my <left> side?	A <table> is on your <left> side.
Q5	Which are the objects I am <close to>?	A <laptop> and a <chair> have been found <very close> to you.
Q6	How many persons have been detected <close to> me?	There are <3> persons <close to> you.
Q7	Are there any <known> persons detected?	<John> is detected <far from> you.

This flexible ontology-based representation allows the system to convey various types of information to the user. In order to deliver semantic content through haptic media, messages generated by the KB should be structured as a *natural language phrase* rather than a simple alert or indication. This information is then communicated to the user by aligning each word or sentence with its haptogram representation, allowing for a semantically richer messaging mechanism.

In the final HIPI prototype, the user will also be able to submit queries through a special panel. For this reason, we have implemented a set of SPIN rules [35] that allow the KBS to respond to queries submitted by the person wearing the HIPI as natural language phrases. Table 2 includes a representative subset of such queries.

IV. INTRODUCING HAPTOGRAMS

A. *From user needs to haptogram design considerations*

Based on an extensive user survey to advance the match between human needs and system support in the specific case of deafblindness, we opted for the transmission of perceptible messages to a grid of actuators mounted onto a smart textile surface on the body. The set of haptic signals had to: (i) be easy to distinguish and perceive; (ii) represent a simple language whose vocabulary is either similar to established vocabularies familiar to the user, or intuitive and simple enough to be learned; (iii) enable sentence building; and (iv) comply with ontology constraints for message generation. Compared with SHC (e.g., in [3] and [4]), where simple and typically environment-related messages are conveyed to the back of the user and other appropriate parts of the body by a human signer, our idea was to similarly convey simple messages, but digitally, by means of actuators. Towards this, a number of complex issues needed to be considered.

A human communicator can simply adapt communication to the actual situational circumstances. For example, in a confined physical space a signer may switch to using the upper arm instead of the back. Similarly, any distortion in reception (as indicated by the body language of the receiver) may be addressed by repetition of the message or further clarification. In an automated system these decisions and interventions need to be preconceived and replaced by other means. Also, in SHC, the signs can easily be adapted and renegotiated between sender and receiver to facilitate perception based on the situational experiences of the receiver. Therefore, next to being idiosyncratic, SHC signs can vary from person to person, and even established tactile sign languages are culturally bound, existing in variants from country to country.

The level of complexity goes beyond this, as by its nature deafblindness involves a very broad spectrum of abilities and impairments, e.g., from those who may have just a handful of concepts in their vocabulary to those who possess a full and rich language ability and can communicate in a sophisticated manner using tactile sign language or even spoken language. This means that the set of concepts and haptogram patterns has to be customizable in due course. Furthermore, SHC involves fine-tuned, compound three-dimensional hand-on-body movements, while the actuator-based haptic patterns need to be rather simple and conveyable in a two-dimensional format, at least for the time being. To handle these research constraints, we fell back on user-centric co-design where potential users' preferences and feedback have been captured. This involved an extensive user study (i.e., 80 detailed interviews conducted in 5 different countries), and ideation workshops (involving participants with deafblindness and communication experts), leading to psychophysics experiments.

The current co-design was constrained by different technological and disciplinary bottlenecks. One of the goals was to build a communication module which could manage at least 100 concepts by haptograms, in the range of the smallest SHC vocabulary known to us [38]. The matrix arrangement of vibrotactile actuators to generate 100 different patterns was considered a viable solution to this end. However, a 3 x 3 actuator grid was too small to allow for this variety, so we opted for a 4 x 4 one as our starting point. On the other hand, static patterns on one's back, not a screen with the best resolution, are more difficult to distinguish than dynamic ones, whereas the latter also extend the number of possible haptograms. Because our ontology at the time of writing this article already included hundreds of concepts to match the capacity of visual detection, 4 x 4 dynamic patterns were a compromise between detection capacity vs. the need to keep things simple for psychophysical experimentation: opting for a larger grid, apart from its wiring problems from a hardware and smart textiles perspective, would have unnecessarily overcomplicated testing. Also, we had to keep in mind that haptogram evaluation by default is an iterative process, to result in designs to be scrapped and replaced time and again.

B. Technical aspects.

Haptograms as a term were proposed by [36]. In their approach, "Haptogram" is a system designed to provide a point-cloud tactile display via acoustic radiation pressure: a tiled 2-D array of ultrasound transducers is used to produce a focal point that is animated to produce arbitrary 2-D and 3-D tactile shapes. The switching speed is very high, so that humans feel the distributed points simultaneously. The Haptogram system comprises a software and a hardware component; the software component enables users to author and/or select a tactile object, create a point-cloud representation, and generate a sequence of focal points to drive the hardware.

Our idea of a haptogram is different and corresponds to logograms. That is, *concepts* are communicated by touch-based drawings, and not character sequences to invoke them. Haptograms are, in short, synthetically generated haptic patterns with a meaning, to be communicated as stimuli to the human body. They could be discretized, i.e., represented by a number of matrix cells, but could also be continuous, illustrated for example as Bezier curves [37]. The meaning of such patterns could be both at word and sentence level. Further, we distinguish between *stable* vs. *changing* patterns and call them *static* vs. *dynamic* haptograms in a communication context, where both can be pulsating for easier recognition. The purpose of haptograms in our framework is to implement an ontology-constrained messaging language for situation awareness assessments, and as raw material for everyday conversation. Both visual and audial information can be translated to haptic patterns.

Such haptograms can be mapped to one or more parts of the body, in single or multiple modes. Unlike studies that focus on using hands (e.g., by developing communicative gloves), our system does not use the hands in order to keep them free for other purposes, such as holding a cane, giving a handshake, examining surfaces or objects, etc. The semantic content is transferred to the body to trigger actuators of vibration, pressure, tapping etc., as well as combinations thereof.

Actuators to display haptograms are advantageously arranged in a matrix form of rows and columns. It is not necessary that they are of equal number; rather the general approach is an *m* x *n* arrangement with *m* rows and *n* columns. A two-dimensional (2D) arrangement of haptic actuators has several benefits over, say, a one-dimensional (1D) linear array arrangement. A single point vibrator - what could be called a zero-dimensional (0D) arrangement - can only convey alarm-like semantic content, such as "Something is happening". (For example, this is the case for cellphones which, while vibrating, can indicate that there is an incoming call or time is out). This way is not best suited to conveying richer content like *what* is happening, or if several things are happening in parallel, including their direction, etc. Compared to 1D, a 2D arrangement enables richer *directional* information: it is possible to identify from where a haptic stimulus is coming. In a 2D arrangement one can use the columns, say, for a certain general class of ontological entities (e.g., related to eating, moving, friends, relationships, etc.) and letting each row to correspond to different concepts therein these realms. This could not be done in 0D and 1D.

The world is inherently three-dimensional but humans have a developed understanding and capability for extracting information out of two-dimensional projections. This holds for exploration both by sight and touch. In the latter case, a 2D vibrator arrangement is much closer to a 2D topological tactile description of a surface than a 1D array. In fact, most often interpreters using SHC describe a room or scene in front of the user by sketching it as a 2D projection, indicating borders, exits and windows, significant objects and persons. A 2D arrangement of actuators lends itself very well to this case. Synonymous is to say that 2D enables *localization* better. 2D arrangements also open up for *dynamic* patterns as actuators corresponding to dots in a haptogram could be made to vibrate in a certain order. This is not only beneficial for perception, compared to the static case, but also yields more possible combinations of patterns over *m* x *n* vibrators. In the limit of increasing *m* and *n* within a given area and even letting *m* and *n* approaching infinity, the resolution increases and continuous patterns are made possible. Of course, human capability to resolve two spatially or temporally separated haptic stimuli is limited, and has to be taken into account. Finally, a 2D arrangement is also a step towards *haptic pictures*, i.e., leaving the paradigm of using symbols. Haptic pictures are such that they mimic a tactile object by levels of contrast of its parts.

Our vocabulary was designed for in-principle receiver testing over a 4 x 4 actuator grid for a proof of concept. We were interested in finding out if the ontology and such a haptic conceptual vocabulary can be aligned, and how pattern sequences reminiscent of sentences can implement the transmission of more complex semantic content. To enable hypothesis testing, we departed from the following:

- Screen on the back, vibrational actuators only:
  - 104 dynamic signs, matching the magnitude of [38];
  - Homonymy disambiguation by different dynamics for the same patterns vs. synonymy representation by two different approaches.
- Vocabulary signs for:
  - Start, stop, alert, sentence type markers, question words, agree/disagree;
  - Known/unknown person, personal pronouns (4);
  - Place adverbs (15), time adverbs (4);
  - Nouns (48), verbs (13), adjectives (3).

Concrete examples below go back to the ontology-compliant sentence sample in Table 2.

C. Examples

As haptograms in principle can be static or dynamic, and can represent both word meaning and sentence meaning, Fig. 5 illustrates the basic idea of the former which was derived from the ASCII code table, where in its matrix cells, instead of characters, concepts are encoded by bit strings in the rows vs. columns.

In Fig. 6, we illustrate two sample dynamic haptograms. In the upper part, numbers indicate the firing sequence of the actuators for concepts (a) and (b), meaning "*stand*" and "*door*". In the lower part the completed shapes of the dynamic haptograms are indicated.

Moving over to sentence meaning, in Fig. 7 we show how three simple statements, "*You are in the kitchen*" (Example 1), "*The bottle is close to the left on the table*" (Example 2), and "*There is an unknown person in the hallway*" (Example 3), can be constructed by concatenating dynamic haptograms. The statement begins with a single-blink sign, indicating the start of a new message, and finishes with a double-blink, pulsating one, indicating end of transmission. Any statement can be accompanied by a separate alert sign to add weight to the communicated content.

Apart from this example, our test included declarative sentences, questions and exclamations to enable a future dialogue between two users with deafblindness, or a user with deafblindness and her/his assistant, family member, educator, etc. Further, the vocabulary is both aligned with the ontology, and is including concepts and parts of speech not covered by the current version, i.e., hints at expansion opportunities. Likewise, e.g., logical operators, numbers, signs for operations etc. can be added following the same line of thought.

Experiment constraints included that over a 4 x 4 grid, no digital numbers could be specified, and screen resolution, i.e., actuator grid granularity was too low for simulating genuine SHC signs.

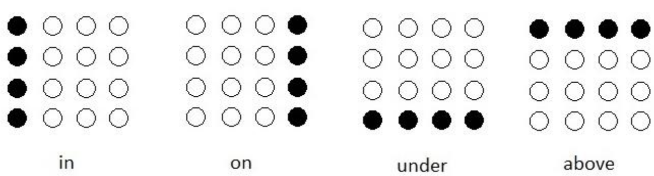


Figure 5. The idea of static haptograms over a 4 x 4 actuator grid.

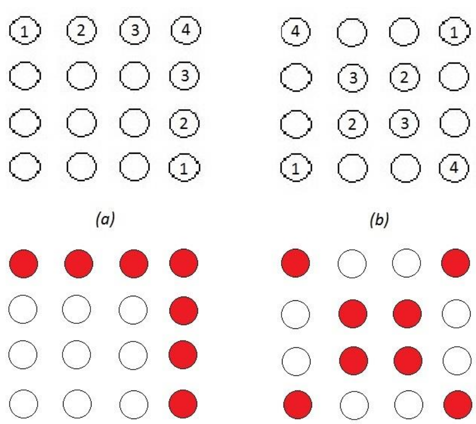


Figure 6. Unfolding sequences of two patterns over a 4 x 4 actuator grid, yielding different dynamic haptograms: (a) "*stand*"; (b) "*door*".

V. Haptogram realization in smart textiles

Haptograms are to be transmitted to different locations on the body and need ample space as they are inherently extended. Very naturally this leads to the use of textiles. Textiles have been there in all human activities throughout the ages, regardless of sex, social status, culture, and occupation. In the form of garments, they cover large parts of the body, such as one's back or upper arms, those areas where haptograms are to be transferred. Important is also that textiles are pliable and adapted to the complicated geometry of the human body. They could be made to tightly fit one's shape, which is very important for any haptic communication as typically, good mechanical contact is central.

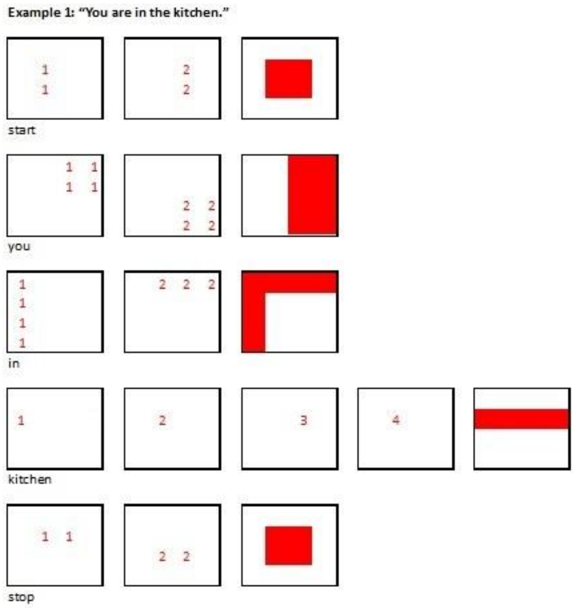
Textile is then a platform. Its enrichment with sensors and actuators, including haptic devices, is often referred to as smart textiles. A basic thought behind the application of haptograms has been to allow the transmission of perceptible messages to the users via actuators mounted into and onto textiles. We opted for discrete haptic signs inherent in the matrix approach above.

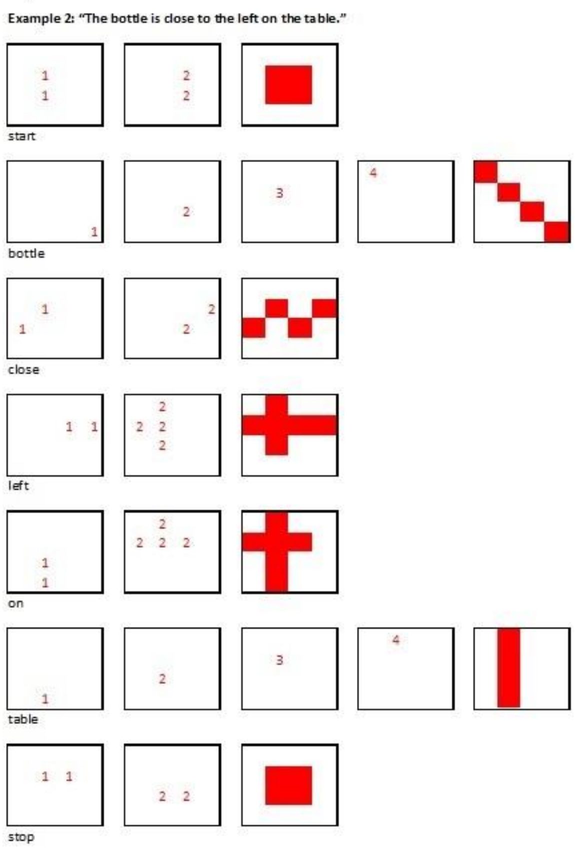
We define *location* as the geometrical position on the body, i.e., an anatomical measure, whereas *placement* is the positioning of some actuator (or sensor) on the textile. A given actuator with a given placement is positioned as close as possible to a target location when taking on the garment, and while wearing it. Further on, we define *cell* as a physical construction on, or in the textile, such as a pocket having a certain placement from which actuation could take place. Typically, we will have many more cells than vibrators in a garment. Cells that have an actuator that is actuating are *active* or firing. Such an active cell corresponds to a *dot* in a *pattern*, a set of active dots. Thus, a pattern of actuators could either simultaneously become active, which is the static version of haptograms, or becoming active in a sequence, i.e., the dynamic alternative.

Last but not least, from a smart textile perspective, a set of cells close to each other constitute a *panel* to host actuators in any particular arrangement. There could be a panel for the back, another for the upper arms, for the waist, and so on.

We constructed a textile testing device for the back (Fig. 8). The human back has a complicated geometry with both convex and concave regions. Added to this is that users are of different sizes. In order to fulfil the simultaneous need of fixed location of actuators, as well as good mechanical attachment for maximal haptic signal transmittance, we opted for a construction with both vertical and horizontal stripes that could be tightened individually at the front by buckles, giving a good fit to the body for all sizes. This structure takes inspiration from a weave with interlaced parts. On the inside, towards the body, detachable pockets are placed, each of them able to host an actuator. These pockets have openings as well so that cables for powering could be taken out and contacted. In total, a very versatile system was created for the realization of haptograms.

In Fig. 9 the placements of 16 vibrotactile elements in a 4 × 4 matrix arrangement is illustrated. The stripes are arranged so that the actuators can be changed from the outside, making it simple to study pattern arrangements without taking off and on the garment.





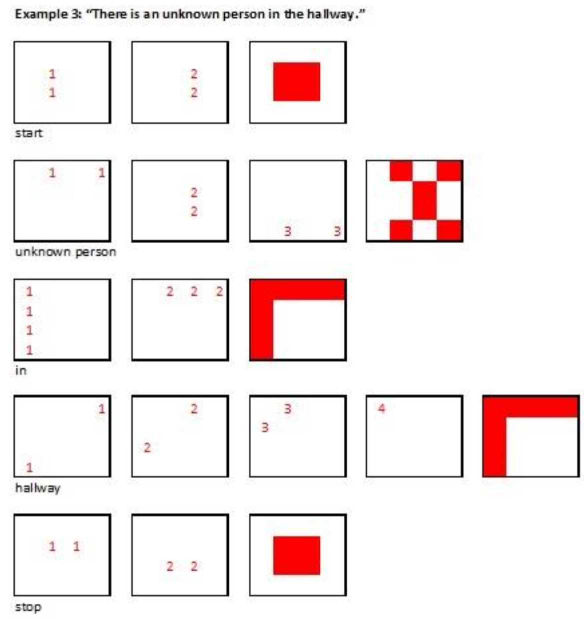


Figure 7. Examples 1-3 are system responses to predefined user queries over a 4 x 4 actuator grid, with patterns plus their unfolding sequences in red.

Figure 8. Back side of the textile testing device. Black squares belong to stripes in the vertical direction, white in the horizontal direction. Together a chessboard like appearance is formed. This results in an inherent matrix for specifying sign component positions.

VI. Conclusion and Future Work

A constraint on this report has been the COVID-19 pandemic which, because of the extreme vulnerability of our user group, prevented the evaluation phase of communication by haptograms and smart textiles. This will be remedied in a next article in due course. Thereby our design considerations had to remain on a theoretical and methodological level.

By means of a proof-of-concept prototype which is work in progress, we adopted a multidisciplinary approach to develop a smart-textile-based communication system for use by people with deafblindness. Focusing on the communication module here, a new ontology connects computer vision with the user to label detected semantic content about objects, persons and situations for navigation and situational awareness. Such labelled content is then translated to a haptogram vocabulary with static vs. dynamic patterns, which are mapped to the body. A haptogram denotes a tactile symbol composed over a touchscreen, its dynamic nature referring to the act of writing or drawing. A vest made of smart textile, in the current variant equipped with a 4 x 4 grid of vibrotactile actuators, was used to transmit haptograms on the user's back. Thereby system messages of different complexity -- both alerts and short sentences -- can be received by the user, who then has the option to respond by pre-coded questions and messages.

The core target groups are users with deafblindness with a diverse range of communication skills, capabilities and needs. The design of our solution is therefore intended to cater for this diversity through personalization, both in terms of the complexity and number of haptograms, and in terms of fit and comfort. The current actuator arrangement made the transmission of 104 dynamic signs possible, including homonyms. This is in the range of the simplest available SHC resource. At the same time, over a 4 x 4 grid, no digital numbers or an alphabet could be specified, and screen resolution was too low to emulate genuine SHC signs.

By means of grids with more actuators, displays with higher resolution can be implemented and tested, paving the way for an extended haptogram vocabulary which can cover more detailed ontology content and, possibly, bridge language barriers among this demographic. Further, we plan to add a mobile sender unit to the prototype to enable de facto two-way communication, and aim to test multi-display message transfer by replacing the smart vest by a smart garment. Psychophysical experiments to evaluate the prototype are in place and in progress.

Acknowledgments

The authors are grateful to Sarah Woodin (University of Leeds) and Astrid Kappers (Eindhoven University of Technology) for helpful suggestions. This paper benefitted from generous advice by four unknown reviewers. This work has been partially funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 780814 SUITCEYES.

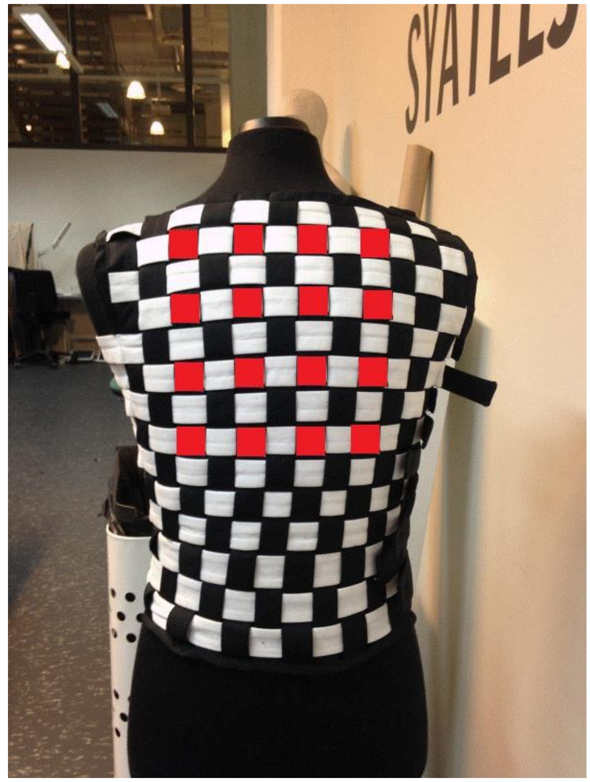


Figure 9. In-principle arrangement of 4 × 4 actuators on the back of the smart vest. Actuators are to be placed beneath the red markers, towards the skin. The non-flat geometry of the human body that any haptic device has to handle is clearly visible.

References

[1] S. Darányi, N. Olson, M. Riga, E. Kontopoulos, and I. Kompatsiaris, "Static and dynamic haptograms to communicate semantic content: towards enabling face-to-face communication for people with deafblindness," paper presented at the "SyMpATHY: SeMAntic Technologies for Healthcare and Accessibility Applications" Workshop at SEMAPRO-19. Porto, 24 September 2019.

[2] A. Israr and I. Poupyrev, "Tactile Brush: Drawing on Skin with a Tactile Grid Display," Proc. SIGCHI Conference on Human Factors in Computing Systems, ACM, pp. 2019-2028, 2011.

[3] R. M. Lahtinen, Haptices and Haptemes: A Case Study of Developmental Process in Social-Haptic Communication of Acquired Deafblind People. PhD dissertation, University of Helsinki, 2008.

[4] R. M. Lahtinen, R. Palmer, and M. Lahtinen, Environmental Description: For Visually and Dual Sensory Impaired People. A1 Management UK, 2010.

[5] J. Chen, R. Turcott, P. Castillo, W. Setiawan, F. Lau, et al., "Learning to Feel Words: A Comparison of Learning Approaches to Acquire Haptic Words," Proc. 15th ACM Symposium on Applied Perception, ACM, pp. 11-18, 2018.

[6] C.M. Reed, H.Z. Tan, Z.D. Perez, E.C. Wilson, F.M. Severgnini, et al., "A phonemic-based tactile display for speech communication," IEEE Transactions on Haptics, 12(1), pp. 2-17, 2018.

[7] M. Walker and K. B. Reed, "Tactile Morse code using locational stimulus identification," IEEE Transactions on Haptics, 11(1), pp. 151-155, 2017.

[8] L.A. Jones and K. Ray, "Localization and Pattern Recognition with Tactile Displays", in Proc. IEEE Symposium on Haptic

Interfaces for Virtual Environments and Teleoperator Systems, 13-14 March 2008, Reno, Nevada, USA, pp. 33-39.

[9] L.A. Jones, J. Kunkel, and E. Piateski, “Vibrotactile pattern recognition on the arm and back,” Perception, vol. 38, pp. 52-68, 2009.

[10] T. L. McDaniel, Somatic ABC’s: A Theoretical Framework for Designing, Developing and Evaluating the Building Blocks of Touch-based Information Delivery. PhD dissertation, Arizona State University, 2012.

[11] S. Kursun-Bahadir, F. Kalaoglu, and V. Koncar, “Analysis of vibrotactile perception via e-textile structure using fuzzy logic.” Fibres & Textiles in Eastern Europe, pp. 91-97, 2012.

[12] P. Lemmens, F. Crompvoets, and J. van den Eerenbeemd, “Design of a tactile jacket with 64 actuators,” Guidelines for Haptic Lo-Fi Prototyping, pp. 21-22, 2004.

[13] R. W. Lindeman, R. Y. Yanagida, and J. L. Sibert, “Towards Full-body Haptic Feedback: The Design and Deployment of a Spatialized Vibrotactile Feedback System,” in Proc. of the ACM Symposium on Virtual Reality Software and Technology (VRST ’04). Association for Computing Machinery, New York, NY, USA, pp. 146–149, 2004, doi:10.1145/1077534.1077562.

[14] F. Gemperle, N. Ota, and D. Siewiorek, "Design of a Wearable Tactile Display," Proc. Fifth International Symposium on Wearable Computers, Zurich, Switzerland, pp. 5-12, 2001, doi:10.1109/ISWC.2001.962082.

[15] U. Yang, Y. Jang, and G.J. Kim, “Designing a Vibro-Tactile Wear for Close Range Interaction for VR-based Motion Training,” in Proc. International Conference on Artificial Reality and Telexistence, pp. 4-9, 2002.

[16] L. Gasparri and D. Marconi, "Word Meaning", *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition), Edward N. Zalta (ed.). Available from: https://plato.stanford.edu/archives/fall2019/entries/word-meaning/ 2020.06.03.

[17] A. Wierzbicka, Semantics. Primes and Universals. New York, Oxford University Press, 1996.

[18] A. Wierzbicka, Understanding Cultures Through Their Key Words. English, Russian, Polish, German, and Japanese. New York, Oxford University Press, 1997.

[19] J. A. Fodor, The Language of Thought. Cambridge, Mass, Harvard University Press, 1975.

[20] N.S. Trubetzkoy, Grundzüge der Phonologie (Travaux du Cercle Linguistique de Prague, 7). Prague, 1939.

[21] Z. Harris, “Distributional structure,” in Papers in Structural and Transformational Linguistics, pp. 775-794, New York, Humanities Press, 1970.

[22] J.R. Firth, “A synopsis of linguistic theory 1930-1955,” Studies in Linguistic Analysis, pp. 1-32, 1957. [Reprinted in F.R. Palmer, Ed. Selected Papers of J.R. Firth 1952-1959. London, Longman, 1968].

[23] L. Wittgenstein, Philosophical Investigations. Oxford, Blackwell, 1967.

[24] R. Carnap, The Logical Syntax of Language. New York, Humanities Press, 1937.

[25] S. Zucchi, Formal Semantics of Sign Languages. Language and Linguistics Compass 6/11, pp. 719-734, 2012.

[26] S. Ojala, Towards an Integrative Information Society: Studies on Individuality in Speech and Sign. TUCS Dissertations No 135, Turku, Turku Centre for Computer Science, 2011.

[27] N. Caporusso, Issues, Challenges and Practices in Advancing Pervasive Human-Computer Interaction for People with Combined Hearing and Vision Impairments. PhD dissertation. IMT Institute for Advanced Studies, Lucca, 2012.

[28] T. Edwards, Language Emergence in the Seattle Deafblind Community. PhD dissertation, Berkeley University of California, 2014.

[29] G. Meditskos and I. Kompatsiaris, “iKnow: Ontology-driven situational awareness for the recognition of activities of daily living,” Pervasive and Mobile Computing, vol. 40, pp. 17-41, 2017.

[30] G. Meditskos, E. Kontopoulos, and I. Kompatsiaris, “Knowledge-driven Activity Recognition and Segmentation Using Context Connections,” in Proc. International Semantic Web Conference, Springer, Cham, pp. 260-275, October 2014.

[31] K. Janowicz, A. Haller, S. J. Cox, D. Le Phuoc, and M. Lefrançois, “SOSA: A lightweight ontology for sensors, observations, samples, and actuators,” Journal of Web Semantics, Vol. 56, pp. 1-10, 2019.

[32] D. Brickley and L. Miller, FOAF Vocabulary Specification 0.99. Namespace Document. [Online]. Available from: http://xmlns. com/foaf/spec/ 2019.08.13.

[33] M. Lefranois, Planned ETSI SAREF Extensions based on the W3C&OGC SOSA/SSN-compatible SEAS Ontology Patterns. Workshop on Semantic Interoperability and Standardization in the IoT, Amsterdam, Netherlands SIS-IoT, 2017.

[34] R. Falco, A. Gangemi, S. Peroni, D. Shotton, and F. Vitali, “Modelling OWL Ontologies with Graffoo,” Proc. European Semantic Web Conference, Springer, Cham, pp. 320-325, 2014.

[35] H. Knoblauch, Spin-SPARQL Syntax. Member Submission, W3C, 2011.

[36] G. Korres and M. Eid, “Haptogram: Ultrasonic point-cloud tactile stimulation,” IEEE Access, Vol.4, pp. 7758-7769, 2016.

[37] D. Salomon, The Computer Graphics Manual. 1st ed., London, Springer, 2011, doi:10.1007/978-0-85729-886-7.

[38] Danish Association of the Deafblind, 103 Haptic Signals – a reference book, 2012, ISBN 978-87-989299-4-9. Available from: http://wasli.org/wp-content/uploads/2013/07/103-Haptic-Signals-English.pdf/ 2020.06.03.

Domain Specific Knowledge Graph Embedding for Analogical Link Discovery

Nada Mimouni

Cedric-Lab, CNAM Paris
Conservatoire National des Arts et Métiers Paris
France
Email: nada.mimouni@cnam.fr

Jean-Claude Moissinac, Anh Tuan Vu

LTCI, Télécom Paris
Institut Polytechnique de Paris
France
Email: jean-claude.moissinac,anh.vu @telecom-paris.fr

***Abstract*—General purpose knowledge bases such as DBpedia and Wikidata are valuable resources for various AI tasks. They describe real-world facts as entities and relations between them and they are typically incomplete. Knowledge base completion refers to the task of adding new missing links between entities to build new triples. In this work, we propose an approach for discovering implicit triples using observed ones in the incomplete graph leveraging analogy structures deducted from a knowledge graph embedding model. We use a neural language modelling approach where semantic regularities between words are preserved, which we adapt to entities and relations. We consider domain specific views from large input graphs as the basis for the training, which we call context graphs, as a reduced and meaningful context for a set of entities from a given domain. Results show that analogical inferences in the projected vector space is relevant to a link prediction task in domain knowledge bases.**

Keywords–Domain knowledge base; Context graph; Entity embedding; Neural language model; Analogy structure; Facts discovery.

I. Introduction

Mining graph structures with machine learning proved to be of great aid to extract hidden useful information in various domains. Recently, learning with graph embedding to encode unstructured data has attracted a lot of attention in research. This paper extends the work in [1] with particular focus on knowledge graphs.

General purpose Knowledge Bases (KB) such as YAGO, Wikidata, and DBpedia are valuable background resources for various AI tasks, for example, recommendation [2], web search [3], and question answering [4]. However, using these resources brings to light several problems which are mainly due to their substantial size and high incompleteness [5] as a result of the extremely large amount of real world facts to be encoded. Recently, vector-space embedding models for KB completion have been extensively studied for their efficiency and scalability and proven to achieve state-of-the-art link prediction performance [6], [7], [8], [9]. Numerous KB completion approaches have also been employed that aim at predicting whether or not a relationship, which is not in the knowledge graphs (KG), is likely to be correct. An overview of these models with the results for link prediction and triple classification is given in [10]. KG embedding models learn distributed representations for entities and relations, which are represented as low-dimensional dense vectors, or matrices, in continuous vector spaces. These representations are intended to preserve the information in the KG namely interactions between entities such as similarity, relatedness, and neighborhood for different domains.

In this work, we are particularly interested in adapting the language modelling approach proposed by [11], where relational similarities or linguistic regularities between pairs of words are captured. They are represented as translations in the projected vector space where similar words appear close to each other and allow for arithmetic operations on vectors of relations between word pairs. For instance, the vector translation

$$v(Germany) - v(Berlin) \approx v(France) - v(Paris)$$

shows relational similarity between countries and capital cities. It highlights the clear-cut analogical properties between the embedded words expressed by the analogy "$Berlin$ is to $Germany$ as $Paris$ is to $France$". We propose to apply this property to entities and relations in KGs as represented by diagrams (a) and (b) in Figure 1. The vector translation example is likely to capture the $capital$ relationship that we could represent by a translation vector $v(capital)$ verifying the following compositionnality [11]:

$$v(France) + v(capital) - v(Paris) \approx 0$$

We use the analogical property for KB completion and show that it is particularly relevant for this task. Our intuition is illustrated by diagrams (b) and (c) in Figure 1, where an unobserved triple can be inferred by mirroring its counterpart in the parallelogram. To the best of our knowledge, leveraging the analogy structure of linguistic regularities for KB completion has never been investigated prior to this work.

We consider applying such properties on excerpts from large KGs, we call context graphs, guided by representative entities of a given domain, where interactions between entities are more significant. Context graphs show to be the bearers of meaning for the considered domain and easier to handle because of their reduced size compared to source graphs.

This paper is organized as follows. Section II gives an overview of related work while Section III presents the global approach. Section IV recalls basic notions of the neural linguistic model. Then, Section V describes our approach to build context graphs and learn features for link prediction. Section VI details the evaluation protocol with expanded algorithms and discusses the initial results and finally, in Section VII we conclude.

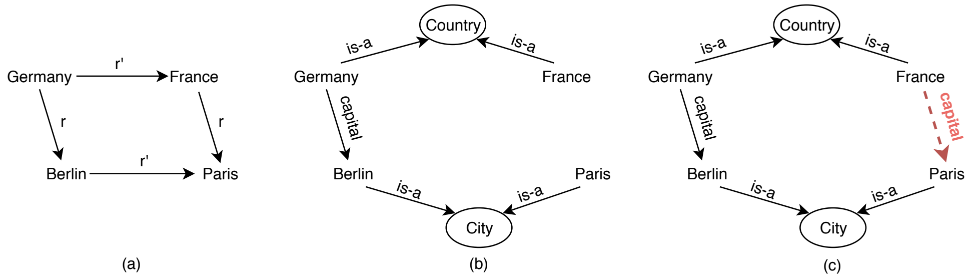


Figure 1. (a) Analogy relation diagram (parallelogram) between countries and capital cities. In KGs (b) and (c), r corresponds to the relation $capital$ and r' is decomposed into two type relations (*is-a*) to concepts $Country$ and $City$.

II. Related work

In what follows we give an overview of research work related to the embedding of knowledge graphs with a particular focus on approaches based on neural language models for the embedding of RDF graphs from which our approach is derived.

A. Knowledge Graph Embeddings

Several KBs exist such as YAGO [12], DBpedia [13], Freebase [14], or WordNet [15]. These knowledge bases contain billions of real-world facts and are, by nature, highly incomplete [5]. Many NLP tasks are based on these resources. The results of word sense disambiguation [16] and question answering [17], [4], for example, are directly impacted by the quality of those graphs.

Vector-space embedding models have a particularly advantageous performance among other statistical relational learning methods [18] proposed for knowledge graph completion or link prediction task [6]. Their main objective is to generalize the graph representation by calculating highly reduced dimensional vectors of entities and their relations. The first proposed algorithm, TransE [7], was inspired by a linguistic embedding model [11] where words are represented as vectors in an embedding space such that linguistic regularities present in the text reflecting relational similarities are captured and modeled as translations in the projected vector space.

Several variations have been subsequently proposed, for example, translation-based models (TransH [19], TransR [20]), bilinear models (DistMult [21], and ComplEx [22]), tensor factorization models (Rescal [23]), neural tensor networks (Ntn [24], [25]) among others [6], [26].

Formally, consider $\mathcal{E}$ to be the set of entities and $\mathcal{R}$ the set of their relations. A knowledge base $\mathcal{K}$ consists of a set of triples $(s, r, o) \in \mathcal{K}$ such that $s, o \in \mathcal{E}$, $r \in \mathcal{R}$, r is the relation between the subject s and the object o.

Each relation r is formulated as a linear map that transforms the subject s, represented as an embedding vector v_r, from its original position near the object o in the vector space. A score function $f(s, r, o)$ is defined by the embedding models to measure the uncertainty of each triple. For example, the TransE f function is as follows:

$$\|v_s + v_r - v_o\|_{l_{1/2}}; v_r \in \mathbb{R}^k$$

where the entities s and o are represented by vectors v_s and $v_o \in \mathbb{R}^k$; k is the vector's dimension. A general margin-based objective function is optimized with different optimization algorithms (e.g., the stochastic gradient descent, abbreviated SGD) to learn the following model parameters: entities vectors and relations vectors or matrices. This function is as follows:

$$\mathcal{L} = \sum_{\substack{(s,r,o)\in\mathcal{K} \\ (s',r,o')\in\mathcal{K}'_{(s,r,o)}}} [\gamma + f(s,r,o) - f(s',r,o')]_+$$

where $[y]_+ = max(0, y)$, γ is the margin hyperparameter; the set of incorrect triples is $\mathcal{K}'_{(s,r,o)}$ generated by altering the correct triple $(s, r, o) \in \mathcal{K}$.

The model Analogy presented in [8] has a particular focus on the study of analogical relations between triples. The approach proposes a formal solution to the problem of multi-relational embedding from an analogical inference point of view by defining analogical properties for entities and relation embeddings as well as optimizing algorithms (objective function) with respect to those properties. The authors argue that analogical inference is particularly advantageous for knowledge base completion and show that their framework improves state-of-the-art performances on benchmark datasets (e.g., WN18 and FB15K).

This assertion supports our decision to exploit the linguistic regularities that reveal analogies between words in order to apply them to entities and their relations in a graph. Since we are working on a particular application domain, we do not apply our method on benchmark datasets. We extract our data from large graphs to target only useful knowledge for this domain. Moreover, our method has the advantage of being based on a simple and efficient model, which directly offers a modeling of the analogy and optimally scales up without the need to define a complex framework for the study of regularities.

B. Neural Language Based RDF Graph Embedding

Another family of techniques for RDF graph embedding uses neural language models and paths in the graph to calculate vectors of entities and relations.

A general technique called Node2vec is proposed in [27]. It aims at creating embeddings for nodes in an (un)directed (a)cyclic (un)weighted graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathcal{W})$ where $\mathcal{V}$ is the set of vertices and $\mathcal{E}$ the set of edges with weights $\mathcal{W}$. The embeddings are learned using the Skip-gram model [28] trained on a corpus of sequences of nodes generated using the sampling strategy. The input graph is turned into a set of directed acyclic sub-graphs with a maximum out degree of 1 using two hyper-parameters for sampling: Return R (probability to go back to the previous node) and Inout Q (probability to explore new parts of the graph).

A closely related approach to our work is described in [29]. The RDF2vec approach uses the neural language model to generate embedding on entities from walks on two general knowledge bases namely DBpedia and Wikidata. Short random walks are created for the whole set of entities in an image of the KB at a given date. Walks using RDF graph kernels are also used on small test datasets due to scalability limitation. The trained models are made available for reuse. The approach we propose here differs in several ways. First, we consider undirected labelled edges in the RDF graph to adapt the neural language model that is compared to directed graph. Second, we use biased walks guided by the application domain to generate sequences that are compared to random walks. Third, rather than using object properties to build the sequences, we consider dataType properties and literals because we assume that they hold useful information for our application domain (e.g., dates, textual descriptions). Finally, and most importantly, we propose to handle scalability issues by contextualizing the input graphs assuming that more relevant information is centralized within a perimeter of α hops around our main entities (α is defined later).

III. KG Completion with Neural Language Model

In the following sections we present the approach we propose for link discovery in KGs. To this aim we use neural language models, namely CBOW and Skip-gram models, that we adapt to the embedding of entities from the KG. Our approach leverages analogical structures extracted from relational similarities between entities generated by the used neural language model. We show in the following how the analogical regularities, applied to entities from KGs, could be used to infer new unobserved triples from the observed ones to complete the original graph. We propose to work on contextualized RDF graphs focused on a specific domain that we extract from large general purpose KGs. We developed a set of algorithms for the construction of context graphs and for the preparation of test data and the evaluation process, which will be detailed as the approach stages progress.

IV. Preliminaries: Neural Language Model

Classic NLP systems and techniques treat words as atomic units represented by indices in a vocabulary without considering similarities between them. Texts are represented as a bag of words using binary feature vectors where each word corresponds to a vector index. Although being simple and robust, such techniques have limited performances in different NLP tasks due to the high dimensional and sparse data vectors generated. With recent advances in machine learning techniques, neural language models have been proposed to overcome these limitations by generating low dimensional distributed representations of words using neural networks. The main goal of word embedding (mapping from words to vectors of real numbers) approaches is to capture as much of the semantic and morphological information as possible from large amounts of unstructured text data. They explicitly model the assumption that words are statistically more dependent as they appear closer in the corpus.

Many different types of models where proposed; however, the earliest suffered from inefficient training of the neural network as they become computationally very expensive on large data sets [30], [31], [32]. An efficient neural model has been proposed in [11], [28], the Word2vec model, and gained wide popularity due to its simplicity. Word2vec is a two-layer neural net model for learning distributed representations of words while minimizing computational complexity. Two different model architectures for parameter learning are used, the Continuous Bag-of-Words model (CBOW) and the Skip-gram model.

A. Continuous Bag-of-Words Model

The intuition behind the CBOW model is the fact of predicting one word while considering a multi-word context within a given window in order to preserve information about the relation of the target word to other words from the corpus. The architecture is shown in Figure 2.

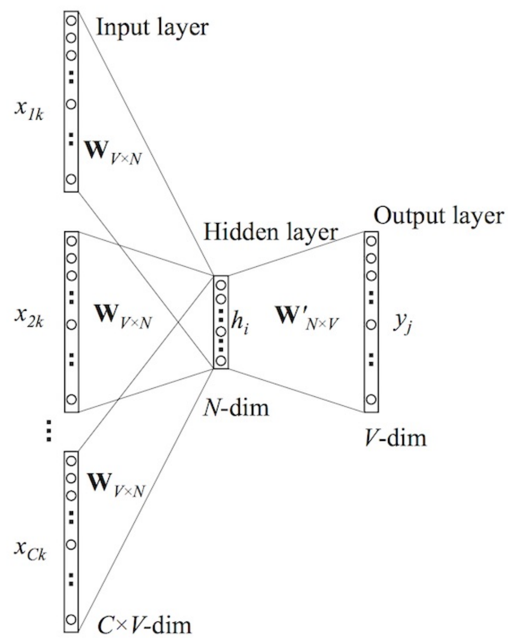


Figure 2. CBOW architecture

The input layer is composed of surrounding words within a given context window c. Their input vectors are calculated from the input weight matrix, averaged and then projected in the hidden layer (or projection layer). The output weight matrix serve to calculate a score for each word in the vocabulary as the probability of being a target word. The architecture of the neural network is formally given by:

- a set of vectors x_i representing training words $w_1, w_2, w_3, ..., w_C$ in the input layer;
- weights matrix W of size $V \times N$ from input to a hidden layer where V is here the entire words vocabulary, and N is the dimension of the hidden layer;
- weights matrix W' of size $N \times V$ as from a hidden to output layer;
- a softmax function as a final activation step.

The goal is to calculate the probability distribution $p(w_i|w_{i-c}...w_{i+c})$, the vector representation of the word with index i, using the softmax function:

$$p(w_i|w_{i-c}...w_{i+c}) = \frac{exp(\overline{v}^T v'_{w_i})}{\sum_{w=1}^{V} exp(\overline{v}^T v'_w)}$$

where V is the entire words vocabulary, v'_w is the output vector of the word w and $\overline{v}$ is the averaged input vector of context words:

$$\overline{v} = \frac{1}{2c} \sum_{-c \leq j \leq c, j \neq 0} v_{w_{i+j}}$$

The objective of the CBOW model is then to maximize the average log probability:

$$\frac{1}{C} \sum_{i=1}^{C} \log p(w_i|w_{i-c}...w_{i+c})$$

B. *Skip-gram Model*

The Skip-gram Model is the opposite of the CBOW model where it is matter to predict c context words from one target word in the input. The architecture is shown in Figure 3.

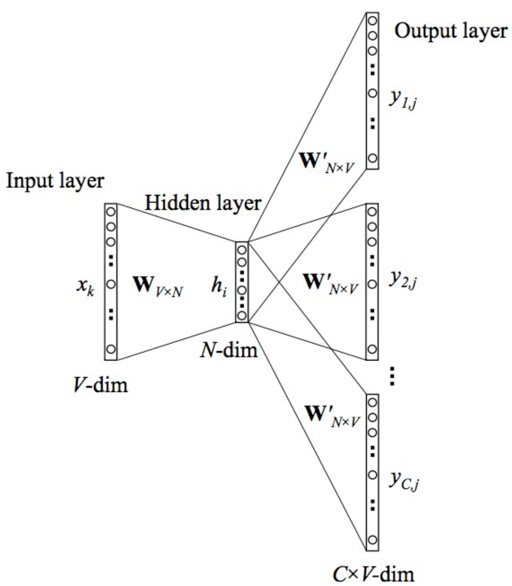


Figure 3. Skip-gram architecture

Formally, given a sequence of training words $w_1, w_2, w_3, ..., w_C$, the objective function to maximize is the following average log probability:

$$\frac{1}{C} \sum_{i=1}^{C} \sum_{-c \leq j \leq c, j \neq 0} \log p(w_{i+j}|w_i)$$

where $-c$ and c are the limits of the context window, word w_i is every word from the working corpus. The first step is then to obtain the hidden layer as:

$$h = W^T x := v_{W_I}^T$$

On the output layer, c multinomial distribution is computed sharing the same weights between the output panels from the hidden to output layer weights matrix W'. The activation of the output use the softmax function to calculate the probability $p(w_{i+j}|w_i)$ as follows:

$$p(w_{i+j}|w_i) = \frac{exp(v'^T_{w_{i+c}} v_{w_i})}{\sum_{v=1}^{V} exp(v'^T_{w_v} v_{w_i})}$$

where V is the entire words vocabulary, v_w is the input vector and v'_w is the output vector of the word w. More details about these techniques can be found in [33], [34].

So far, the models presented above, both CBOW and Skip-gram, are in their original forms without any efficiency optimization applied. Computed in a straightforward manner, the computational complexity of those algorithms corresponds to the size of the vocabulary. In fact there exist two vector representations for each word in the vocabulary, v_w and v'_w, the input and the output vector, respectively. During the update process it is required to iterate through every word w_j in the vocabulary, compute their probability prediction, and their prediction error to finally update their output vector v'_j. Doing these computations for all of the words for every training instance makes the learning of output vectors considerably expensive and the whole learning process impractical to scale up for large vocabularies. Two optimization techniques were proposed to solve this problem by limiting the number of output vectors that require an update per training instance, hierarchical softmax and negative sampling [11].

As stated by the authors in [11], negative sampling outperforms hierarchical softmax on the analogical reasoning task, which is our main concern in this paper. They also argue that the linearity of the Skip-gram model makes its vectors more suitable for such linear analogical reasoning. This will guide our parameters calibration for training the word2vec model in the evaluation step.

V. APPROACH

The approach we propose in this work adapts the neural language model word2vec for the embedding of knowledge graphs. It benefits from the properties offered by this model, which explicitly assume that closer words in text documents are statistically more dependent. For RDF graphs, words are replaced by entities and relations between entities. Sequences of entities and relations should be generated from the RDF data in order to apply the word2vec model as on sentences of words. After graph conversion, we can train the neural language model to represent entities and relations between them as vectors of numerical values in a latent feature space. These vectors can directly be used to perform analogical inferences for link prediction. In our approach, we work on excerpts from large knowledge graphs, we call context graphs, which represent condensed information from a given domain. Many reasons motivate our choice; we discuss them in the following subsection and show how such context graphs are built and explored.

A. *Building Context Graphs*

For the remainder of this article, we will study a knowledge graph extract; we name it Context Graph (CG). We first start by explaining why CGs and, more generally, extracts from knowledge graphs, are interesting to study and how they can be constructed.

1) Motivation For Context Graphs: In the semantic web community, the concern has arisen to work with extracts from large graphs, rather than with large graphs themselves. Various motivations are encountered for this. To begin with, a query on a graph, in particular a `CONSTRUCT` query, creates a 'view' on this graph that constitutes an extract on which we generally wish to do targeted operations: insertion into another graph, display, processing, etc. We may want to put aside the obtained extract, for example, because the corresponding request is long to process and therefore costly in performance for the triple store of the original graph.

An extract targeted on a class or a category of objects can be useful, for example, for an analysis and an improvement of the quality of these types of objects in the large graph: used properties, missing properties, consistency with a schema, etc.

We can assume that queries on an extract can execute faster than the same queries on the full graph (and give the same results for a category of queries targeted by the extract). This can be a way to bypass timeout problems on graphs such as DBpedia. We have no precise benchmarks on this subject. To create such benchmarks; it is necessary to install a copy of DBpedia on the same computer, then build the extract, and test a list of queries, which is out of scope in this work.

The same is true for other processings on graphs. For example, in [29], building a model from DBpedia may take up to several days with a powerful machine [35] while building a model for this work takes 11 minutes on a laptop with four GB of memory and a quad-core 2.8 GHz processor (see more details at the end of the evaluation section). We argue that with smaller and more focused graphs as our Context Graphs, it is easier and more reliable to apply inference methods.

Extracts can also be shared more easily to contribute to students work. A dated and versioned excerpt can also be a significant dataset to support research work by having access to the data that have been used for this work. This is particularly useful if you are working on constantly evolving data graphs such as Wikidata.

2) Context Graphs Building Process: We define a Context Graph as a sub-graph of a general $\mathcal{KG}$ (e.g., DBpedia) representative of a domain D.

1) The first step to build a CG is to identify a list of seeds defining the domain. A seed is an entity from $\mathcal{KG}$ corresponding to a concept that is considered relevant for D. For example, if the concept is '*Musée du Louvre*', the corresponding entity in DBpedia is <http://dbpedia.org/resource/Louvre>. In some domains this list is obvious as for museums, hotels, or restaurants. In general, the common practice is to rely on a reference dataset (such as IMDB for cinema).
2) The second step extracts from $\mathcal{KG}$, the neighborhood for each seed within a given depth filtering useless entities or predicates (not informative for D) and returns the final CG as the union of elementary contexts. We use the CG in the following algorithm as the basis for the embedding model. For example, DBpedia-fr uses the predicate <http://dbpedia.org/ontology/wikiPageRevisionID>, which is maintenance information about the source of the entity. The node revision ID is inadequate for our work and thus we consider it as out of our target domain.

Algorithm 1: *CONTEXT BUILDER*

```
Function ContextBuilder(KG, seedsEntities, radius, filteredEntities)
    Input  : A knowledge graph KG
             A neighborhood depth to reach radius
             A set of entities which are used as seeds seedsEntities
             A set of entities which are excluded from the seeds filteredEntities
    Output : Context Graph context
    level ← 0
    context ← ∅
    while level < radius do
        newSeeds ← ∅
        foreach s ∈ seedsEntities do
            C_s ← FindNeighbors(KG, s)
            context ← context ∪ C_s
            newSeeds ← newSeeds ∪
              EntityFilter(C_s, filteredEntities)
        end
        level ← level +1
        seedsEntities ← newSeeds
    end
    context ← context ∪ AddClasses(KG,
      Entities(context))
    return context
```

We create the algorithm `CONTEXT BUILDER` (Algorithm 1) to build a context graph $context$ from a knowledge graph $\mathcal{KG}$ for a given domain D. For a set of seeds ($seedsEntities$), `findNeighbors`(s) extracts a neighboring context C_s from a knowledge graph $\mathcal{KG}$ for each seed s. The final context, $context$, is updated adding C_s. A list of new seeds, $newSeeds$, is updated with the new collected entities after filtering the terminal nodes with the `EntityFilter` method. The exploration depth $level$ is incremented by 1 at each step up to the desired $radius$ limit. At the end of process, the resulting context $context$ is expanded with the classes of entities extracted from $\mathcal{KG}$ by the methods `AddClasses` and `Entities`.

The purpose of the `EntityFilter` method is to eliminate a certain number of entities that seem useless for a given application. The entities to be filtered are given in the list *filteredEntities*. In our application, examples of elements in this list are entities belonging to the T-Box due to their very general nature (e.g., in DBpedia: dbo:Building, dbo:Place or owl:Thing) and structuring nodes (e.g., DBpedia-fr pages layout, <http://fr.dbpedia.org/resource/Modèle:P.>). Those nodes introduce a great deal of noise without bringing relevant information for the targeted domain. For example, in our experiments on DBpedia, we found $3,486$ entities built on <http://fr.dbpedia.org/resource/Modele:????> giving rise to $2,692,515$ links and $101,235$ links to <http://www.w3.org/2004/02/skos/core#Concept>.

The `Entities` method finds all the entities in the created context graph. The `AddClasses` method adds the classes of all the entities after finding them in the source graph. At this stage, we also add the ontology (the T-Box) that organizes these classes, if it exists. Thus, we have a knowledge of the nature of the manipulated entities, which allows us to draw conclusions on the basis of these classes as we will see below. It should be noted that the concept of a class can differ from one knowledge graph to another; this is, for example, the case for Wikidata that uses its own concepts, while DBpedia uses the concepts defined by RDFS [36].

B. Context Graph Sub-structures

We transform the entities and relations in the CG as paths that are considered as sequences of words in natural language. To extract context graph sub-structures, we use the breadth-first algorithm to get all the graph walks or random walks for a limited number N.

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be an RDF graph where $\mathcal{V}$ is the set of vertices and $\mathcal{E}$ is the set of directed edges. For each vertex $v_x \in \mathcal{V}$, we generate *all* or N graph walks P_{v_x} of depth d rooted in the vertex v_x by exploring indifferently direct outgoing and incoming edges of v_x. We iteratively explore direct edges, both incoming and outgoing, of neighbors v_{x_i} of vertex v_x until depth d is reached. The paths after the first iteration follow this pattern $v_x \rightarrow e_i \rightarrow v_{x_i}$ where $e_i \in \mathcal{E}$. The first token of each path $p \in P_{v_x}$ is the vertex v_x followed by a sequence of tokens that might be labels of edges or vertices. The final set of sequences for $\mathcal{G}$ is the union of the sequences of all the vertices $\bigcup_{v_x \in \mathcal{V}} P_{v_x}$.

Hereafter are some examples of entity sequences that have been extracted from DBpedia using graph walks of depth 2, 4 and 8, respectively:

dbr:Walker,_Texas_Ranger → dc:subject → dbr:Category:Martial_arts_television_series

dbr:Walker,_Texas_Ranger → dbo:creator → dbr:Paul_Haggis

dbr:Walker,_Texas_Ranger → dbp:country → United States

fdbr:Maison_de_Balzac → dbo:wikiPageWikiLink → fdbr:Honoré_de_Balzac → fdbp:auteur → Patricia Baudouin

fdbr:Maison_de_Balzac → dbo:wikiPageWikiLink → fdbr:La_Comédie_humaine → dbo:wikiPageWikiLink → fdbr:Venise → dbo:wikiPageWikiLink → fdbr:Paul_Bourget → dbo:wikiPageWikiLink → fdbr:Catégorie:Grand_officier_de_la_Légion_d'honneur

The following correspondences are used for properties:

fdbr	http://fr.dbpedia.org/resource/
dbo	http://dbpedia.org/ontology/
dbp	http://dbpedia.org/property/
dc	http://purl.org/dc/terms/

C. Feature Learning

Next, we train the word2vec neural language model, which estimates the likelihood of a sequence of entities and relations appearing in the graph and represents them as vectors of latent numerical features. To do this, we use the continuous bag of words (CBOW) and Skip-gram models as described in Section IV. CBOW predicts target words w_i from context words within a context window c while Skip-gram does the inverse and attempts to predict the context words from the target word. The probabilities $p(w_i|w_{i-c}...w_{i+c})$ and $p(w_{i+j}|w_i)$ are calculated using the softmax function.

Once the training is finished, all entities and relations are projected into a latent feature space where semantically similar entities are positioned close to each other. Moreover, we can perform basic mathematical operations on the vectors in order to extract different relations between entities.

D. Analogical Inference

In this last step, we extract analogical properties from the feature space to estimate the existence of new relationships between entities. We use the following arithmetic operation on the feature vectors (entities of Figure 1): $v(Berlin) - v(Germany) + v(France) = v(x)$ which we consider is solved correctly if $v(x)$ is most similar to $v(Paris)$. On the left-hand side of the equation, entities contribute positively or negatively to the similarity according to the corresponding sign. For example, $Germany$ and $France$ having the same type $Country$ contribute with different signs, $Berlin$, of a different $City$ type, contribute with the opposite sign of the corresponding Country. The right-hand side of the equation contains the missing corner of the diagram, which remains to be predicted. We then use cosine similarity measures between the resulting vector $v(x)$ and vectors of all other entities of the same type in the embedding space (discarding the original ones of the equation) in order to rank the results.

In this article, the assessments focus on relationships between museums and artists. We will therefore be interested in equations on the model $v(artist1) - v(museum1) + v(museum2) = v(x)$. In addition, we verify that x has the good type, here Artist.

E. Semantic filtering

Analogical inference returns as result a set of entities who's vectors are similar to the reference vectors.

It is possible to acquire entities which do not have the required type (City in the example). To handle this, we apply a semantic filtering to keep only entities of the expected type. Doing so, we use a combination of embeddings and semantics to achieve filtered results to explore for evaluation.

VI. Experimental Evaluation

In what follows we present the experimental evaluation of the proposed approach with a case study in the tourism field, particularly the museums of Paris. Data is extracted from general purpose KBs on which the approach is applied and evaluated against ground-truth of artists and museums analogies.

A. Case Study

We test our approach on a sub-graph of DBpedia representing a target domain; here we chose museums of *Paris Musées*, a federation of museums in Paris. We propose to address the scalability issue by contextualizing the input graphs, assuming that more relevant information is centralized within a perimeter of α hops around the main entities of this domain. We used $\alpha = 2$ as suggested by [37]. We build our $\mathcal{KG}$ as the union of individual contextual graphs of all entities representing the input data from the cultural institution *Paris Musées* (12 sites). We identify each site by its URI on DBpedia-fr after an entity resolution task (in the following, we denote the URI http://fr.dbpedia.org/resource/entity shortly as dbr:entity).

Obtaining URIs is a classic Entity Linking problem. As we are only dealing with a small number of elements, we used a rudimentary method: starting from the name of each museum, a search in DBpedia-Fr of entities with the same label - by testing different variants of a character case - we get a set of propositions that are validated by a person. It is also a person who carries out a search in DBpedia-Fr for entities that could not be found by this process. Of course, for a greater number of entities describing a domain, it would be necessary to use an advanced method of Entity Linking and a context for each entity to avoid any ambiguities [38].

The final graph contains $448,309$ entities, $2,285$ relations, and $5,122,879$ triples. To generate sequences of entities and relations we use random graph walks with $N = 1000$ for depth $d = \{4, 8\}$. We also consider for each entity all walks of depth $d = 2$ (direct neighbours).

The context graph for *Paris Musées* was constructed with a depth of 2. The core of the context therefore has a depth of 1. Studying the impact of choosing this depth is beyond the scope of this article. A *blacklist* has been created essentially comprised of all the elements of the *T-Box*, considered as terminal nodes. For example, if a node brought us to `owl:Thing` and we followed the links from there, clearly, we would bring back $1,527,645$ entities that are not necessarily related to our domain.

Table I gives a description of a context graph extracted from DBpedia-fr for the depth $N = 2$. We test different settings for the constitution of such a graph. The values in this table are therefore only indicative.

TABLE I. Description of a context graph $\mathcal{CG}$ extracted from DBPedia-fr with $N = 2$

	Context $\mathcal{CG}$	DBPedia-fr	%
Distinct entities	448309	10515624	4.2
Distinct predicates	2285	20322	11.24
Links	5122879	185404534	2.76
Links by entity (mean)	11.42	17.63	

It is normal that there are fewer links per node in $\mathcal{CG}$ than in DBPedia-fr, since, by construction, we have eliminated some links that are not very informative in our application framework as explained above.

We therefore have a number L of links 36 times lower and a number S of vertices 23 times lower in the $\mathcal{CG}$ that in $\mathcal{KG}$. On an algorithm in $O(L + S)$, such as the Breadth-first search, we can anticipate a gain factor of around 30, which can strongly contribute to the applicability of a large number of methods. The gain can become considerable on algorithms such as those of the shortest path search between two nodes, when it is also a matter of giving weights to edges, the complexity of which is in $O(S^2)$.

We then train the Skip-Gram word2vec model on the corpus of sequences with the following parameters: window size = 5, number of iterations = 10, negative samples = 25 (for the purpose of optimisation), and dimension of the entities' vectors = 200. We use gensim implementation of word2vec [39]. The parameters values are inspired by those used for the training of the RDF2vec model [29] and, for the remaining ones, the default values, as proposed in [11], are used. Moreover, we trained our model with the CBOW method and with a larger vector dimension (500). We notice in general better performance with the Skip-Gram method, but cannot make any assertion about the vector dimension.

Our method cannot be evaluated against others using benchmark datasets such as FB15K, WN18 [7], [8]. It requires defining a context and extracting a subgraph from it; none of the other methods use such a context in the available evaluations and our proposal is strongly linked to the definition of such a context. So far, we have no knowledge of any experiments that rely on DBpedia-Fr. Some rely on DBpedia, but French museums are poorly represented in DBpedia.

B. Evaluation Protocol

Existing benchmarks for testing analogy tasks in the literature are designed for words from text corpora. To the best of our knowledge, using a language model driven analogy for link prediction in knowledge graphs has not been investigated yet.

To evaluate our approach, we build a ground-truth for an analogy between entities in the KG. Each entry corresponds to a parallelogram as described in Figure 1, with one unobserved triple in the $\mathcal{KG}$. For each entity, corresponding to a museum site in our application, we collect a list of well-known artists for this site as follows: find in DBpedia-fr the list of artists (dbo:Artist) or otherwise, individuals (dbo: Person) who are associated with the site. For some sites, we manually create the list, for example, by searching for well-known artists for a museum on the website [40].

The data used for the evaluation is made up of two tables indexed by the museum identifier. The first table associates a museum with an artist considered as key for this museum; for example, Zadkine for the Zadkine museum. The second table associates a list of people - essentially artists - important for this museum. These lists were established by consulting the list of artists from the Paris Musées Collections site [40] and consolidating this information with the corresponding Wikipedia pages. For each artist and museum in the data, we use the URI in DBpedia-Fr.

In the following paragraphs of this subsection, we first present the method used to obtain the results. Then, we propose two ways to evaluate them.

We create the algorithm `FIND SIMILARS` (Algorithm 2) which builds a two-dimensional table. The first dimension is the museum for which we search artists. The second dimension is the museum for which we know a principal artist. As the result, each cell of the table contains a list of proposed artists.

The first test is formalized as follows: given a pair ($museum_a$, $artist_a$ = main artist for $museum_a$), for each $museum_b$, $a \neq b$, we search for a list of artists $artist_b$, which has a similar role as the role of $artist_a$ for $museum_a$ (here, we consider the artist who has the most works exhibited or who is the most present in the museum, that we call the main artist). Then, we assess the pertinence of each $artist_b$ for $museum_b$.

The evaluation test aims at discovering $artist_a$ for $museum_a$ considering a known triple <$museum_b$, $artist_b$> while varying b and measuring the mean of the returned results. We use conventional metrics: Mean Reciprocal Rank (MRR) and the number of correct responses at a fixed rate (Hits@).

Algorithm 2: *FIND SIMILARS*

Function `FFindSimilarArtists`($MainArtists,$ $Museums,$)
  Input :
  List of main artist uris, one by museum $MainArtists$
  List of museum uris $Museums$
  Max of returned similar entities Max
  Output :
  Table of proposed similar artists for M_j given (M_i, A_i) $SimilarArtists$ [][]
  foreach $M_i \in Museums$ **do**
    $A_i \leftarrow MainArtists$ [M_i]
    foreach $M_j \in Museums$ **do**
      $Similars$ [M_i][M_j] $\leftarrow$ `FindSimilars`(A_i, M_i, M_j, Max)
      $SimilarArtists$ [M_i][M_j] $\leftarrow$ `TypeFilter`($Similars$ [M_i][M_j], $tArtist$)
    end
  end
  return $SimilarArtists$

Algorithm 3: *EVALUATION A*

Function `MuseumsFocusedEvaluation`($SimilarArtists,$, $Museums$)
  Input :
  Table, result of Algorithm 2 $SimilarArtists$
  Reference lists of artists, one list by museum
  List of museum $Museums$
  Output :
  List, mean of MRR values by museum $MMrr$
  List, mean of Hit@3 values by museum $MHit3$
  List, mean of Hit@5 values by museum $MHit5$
  List, mean of Hit@10 values by museum $MHit10$
  foreach $M_j \in Museums$ **do**
    foreach $M_i \in Museums$ **do**
      $Proposal \leftarrow SimilarArtists$ [M_i][M_j]
      Mrr [M_i], $Hit3$ [M_i], $Hit5$ [M_i], $Hit10$ [M_i] $\leftarrow$ `EvaluateAgainstRef`($Proposal,$ $ArtistsRefList$ [M_i])
    end
    $MMrr$ [M_j], $MHit3$ [M_j], $MHit5$ [M_j], $MHit10$ [M_j] $\leftarrow$ `Mean`($Mrr,$ $Hit3,$ $Hit5,$ $Hit10$)
  end
  return $MMrr$, $MHit3$, $MHit5$, $MHit10$

The evaluation protocol is as follows: for each URI M_i, URI of a museum, let URI A_i be the URI of the first artist identified for M_i, consider all $M_j \mid j \neq i$, and find the top most similar entities of the predicted vectors with positives = [A_i,M_j] and negative =[M_i]. In the list of results, we filter by type $Artist$, and then examine the intersection with artists A_l associated with M_j.

In other words, we fix the target M_j and look at the obtained results by varying the reference pair (M_i, A_i). We consider the relevance of the returned result on M_j as the average relevance over all pairs (M_i, A_i). Algorithm `EVALUATION A` (Algorithm 3) shows the steps of the first evaluation test.

In Algorithm 3, the `EvaluateAgainstRef` function uses the list of proposed artists and the reference list for museum M_i. This function checks whether the artists expected by the reference list have been found in the 3^{rd} -or 5^{th} or 10^{th}- first returned results, and compares the rank of found results with the reference ranking (with MRR metric). The `Mean` function averages the obtained results for the different M_i tested.

It is worth noticing that we frequently find loosely qualified links between museums and artists; such links are very common in DBpedia and use the property `wikiPageWikiLink` representing an untyped link. Subsequent work is required to qualify them.

C. Results

Table II shows the results of MRR and Hits@$\{3, 5, 10\}$ (%) for $d = \{4, 8\}$ and $N = 1000$. The final row of Table II with columns $d = 8$ shows the impact of considering longer paths on the performances of the approach. In fact, longer paths capture richer contexts for entities and results in better vectors estimation by the neural language model.

We compared our approach with the one presented in [29], which creates a model, modelDB, for all entities in DBpedia. For each entity in our ground-truth built on DBpedia-fr, we look for its equivalent in DBpedia and verify that it is contained in the vocabulary of modelDB built with $d = 4$. Only 7 out of 12 museum entities are in modelDB, as well as their first associated artist among others. The analogy tests return globally poor results. ModelDB were unable to retrieve relevant entities in the top 100 returned answers, as was the case for our model trained on the CG, without any improvement even if extended to top 5000. This result is to be expected when we look at the following table, which shows that our CG has better coverage of the ground-truth domain entities, mainly artists, compared to DBpedia.

TABLE III. GROUND-TRUTH ENTITIES IN DBPEDIA AND DDBPEDIA-FR.

	dbo:Person	dbo:Artist	No type	dbo:Museum
DBpedia	272	190	44	7
DBpedia-fr	272	327	6	12

The first row of Table III shows that not all dbo:Artist are linked to dbo:Person (ex: dbr:Sonia_Delaunay). With 12 museums and 334 artists in the reference list, 97.90% can be identified as an artist in our context graph vs. 56.88% in DBPedia, which partly explains the poor results with modelDB. As we filter the returned results by type Artist (or more generally by Person), several relevant answers are filtered.

We also compared our approach with a random selection of entities of type dbo:Artist in the vocabulary of the model. The results, given in columns $d = 4R$ of Table II, show a great benefit of leveraging the regularities in the vocabulary space to extract relationships between entities.

While analysing values on Table II, we noticed wide discrepancies between the results of different museums. For

TABLE II. MRR AND HITS@$\{3, 5, 10\}$ (%) OF A SUBSET OF REPRESENTATIVE EXAMPLES OF *Paris Musées* DATA FOR $d = \{4, 8\}$ AND $N = 1000$ WITH ANALOGY AND RANDOM FOR $d = 4$ (D=4R).

	MRR			Hits@3			Hits@5			Hits@10		
Entity	d=4R	d=4	d=8	d=4R	d=4	d=8	d=4R	d=4	d=8	d=4R	d=4	d=8
dbr:Musée_Bourdelle	0.05	0.39	0.43	0.09	0.50	0.42	0.18	0.50	0.42	0.18	0.66	0.50
dbr:Musée_Carnavalet	0.01	0.43	0.59	0.00	0.58	0.67	0.09	0.66	0.75	0.09	0.83	0.75
dbr:Musée_Zadkine	0.00	0.43	0.44	0.00	0.41	0.42	0.00	0.50	0.50	0.00	0.50	0.50
dbr:Musée_Cernuschi	0.01	0.42	0.50	0.00	0.50	0.58	0.00	0.58	0.67	0.09	0.75	0.67
dbr:Petit_Palais	0.04	0.38	0.63	0.09	0.50	0.75	0.09	0.66	0.75	0.09	0.66	0.75
dbr:Maison_de_Balzac	0.03	0.23	0.44	0.09	0.25	0.58	0.09	0.41	0.58	0.09	0.41	0.58
dbr:Musée_Cognacq-Jay	0.09	0.33	0.49	0.09	0.33	0.58	0.09	0.33	0.58	0.09	0.33	0.58
dbr:Musée_d'art_moderne	0.04	0.36	0.71	0.09	0.41	0.75	0.09	0.50	0.83	0.09	0.58	0.83
dbr:Musée_Romantique	0.03	0.34	0.48	0.09	0.41	0.50	0.09	0.41	0.58	0.09	0.50	0.58
dbr:Palais_Galliera	0.00	0.36	0.48	0.00	0.50	0.50	0.00	0.50	0.58	0.00	0.50	0.58
dbr:Maison_de_Victor_Hugo	0.01	0.38	0.55	0.00	0.50	0.58	0.00	0.58	0.58	0.18	0.58	0.67
dbr:Musée_de_Grenoble	0.00	0.34	0.33	0.00	0.41	0.33	0.00	0.50	0.33	0.00	0.50	0.33
All entities in *Paris Musées*	0.02	**0.37**	**0.52**	0.04	**0.44**	**0.58**	0.06	**0.51**	**0.62**	0.09	**0.57**	**0.64**

example, Hits@10 values for dbr:Musée_d'art_moderne and dbr:Musée_de_Grenoble are respectively: 0.83 and 0.33. This impacts the global performance of all museums (see last row of Table II). The result means for the second value that the system was not able to retrieve the corresponding artist for dbr:Musée_de_Grenoble in the top returned results. We argue this is mostly related to the representativeness of this artist's entity in the KG and how it is linked to the museum's entity; less interlinked entities (directly or indirectly through neighbors) have a lower chance of being related with the analogy structure in the embedding space.

To explain this, we run the following evaluation test:

- goal: evaluate how well we find museum which has $artist_a$ as main artist,
- input: each known pair ($museum_b$, $artist_b$) where $artist_b$ is the main artist for $museum_b$,
- method: we look for museums that play a similar role for $artist_a$ to which played by $artist_b$ for $museum_b$ and verify if the returned museums have a relationship with the $artist_a$ in the ground-truth

For example, consider the following pair $(museum_b, artist_b) = (Maison\_de\_Victor\_Hugo, Victor\_Hugo)$ and the relation $main\_artist\_for$ between them, and consider $artist_a = Honore\_de\_Balzac$. We want then to find museums related to $Honore\_de\_Balzac$ with the $main\_artist\_for$ relation (as by analogy of the relation between the pair $(Maison\_de\_Victor\_Hugo, Victor\_Hugo)$. In other words, find museums for which $Honore\_de\_Balzac$ is the main artist.

The evaluation protocol is as follows: for each $Auri_i$, URI of an artist, consider all known triples <$Muri_j$, $Auri_j$ > | $j \neq i$, and find the top most similar entities of the predicted vectors ranked by similarity. In the list of results, we filter by type $Museum$, and we then examine the intersection with museums $Muri_l$ associated with $Auri_i$.

In other words, we fix the pair (M_i, A_i) and assess all possible targets M_j in the returned results. We consider here that the pertinence for the artist A_i is the average pertinence for all possible targets. Algorithm `EVALUATION B` (Algorithm 4) shows the steps of this second evaluation.

In Algorithm 4, the `EvaluateAgainstRef` function uses the list of proposed artists and the reference list for the

Algorithm 4: $EVALUATION\ B$

```
Function ArtistsFocusedEvaluation(SimilarArtists, ArtistsRefList, Museums)
    Input :
        Table, result of Algorithm 2 SimilarArtists
        Reference lists of artists, one list by museum
        List of museum Museums
    Output :
        List, mean of MRR values by museum MMrr
        List, mean of Hit@3 values by museum MHit3
        List, mean of Hit@5 values by museum MHit5
        List, mean of Hit@10 values by museum MHit10
    foreach M_i ∈ Museums do
        ← [M_i ] foreach M_j ∈ Museums do
            Proposal ← SimilarArtists [M_i ][M_j ]
            Mrr [M_j ], Hit3 [M_j ], Hit5 [M_j ],
              Hit10 [M_j ] ←
              EvaluateAgainstRef(Proposal,
              ArtistsRefList [M_j ])
        end
        MMrr [], MHit3 [], MHit5 [], MHit10 []
          ← Mean(Mrr, Hit3, Hit5, Hit10)
    end
    return MMrr, MHit3, MHit5, MHit10
```

museum M_j. This function checks whether the artists expected by the reference list have been found in the 3^{rd} -or 5^{th} or 10^{th}- first returned results and compares the rank of found results with the reference ranking (with MRR metric). The `Mean` function averages the obtained results for the different M_j tested.

Table IV shows the results of MRR and Hits@$\{3, 5, 10\}$ (%) for $d = 4$ and $N = 1000$. For example, the line dbr:Antoine_Bourdelle is built for the given pair (Musée Bourdelle, Antoine Bourdelle), then, for each other museum, we search the proposed artists. We evaluate the results with MRR, Hit@3, Hit@5, Hit@10, and put in the table the mean of the results for each museum.

The wide differences between artists' results in the last column of Table IV (Hits@10) (e.g.,dbr:Victor_Hugo and dbr:Geer_Van_Velde) reveals the impact of the triple interlinkage in the graph on the analogy prediction test. Thus, good prediction performance of new triples could be achieved with a good representativeness of known triples by the context graph. For evidently strong interlinkages such as for (Musée Victor Hugo, Victor Hugo), (Musée Bourdelle , Antoine Bourdelle), we have satisfactory results. For a weaker link, such as (Musée Carnavalet, Israël Silvestre), the results are unsatisfactory. We need to investigate what can be foreseen as a good interlinkage.

TABLE IV. MRR AND HITS@$\{3, 5, 10\}$ (%) OF REPRESENTATIVE EXAMPLES OF ARTISTS EXHIBITED IN MUSEUMS OF *Paris Musées* FOR $d = 4$ AND $N = 1000$

Entity	MRR	Hits@3	Hits@5	Hits@10
dbr:Antoine_Bourdelle	0.61	0.72	0.81	0.81
dbr:Israël_Silvestre	0.08	0.09	0.13	0.13
dbr:Gustave_Courbet	0.38	0.45	0.45	0.72
dbr:Ossip_Zadkine	0.67	0.81	0.90	0.91
dbr:Xu_Beihong	0.74	1.0	1.0	1.0
dbr:Honoré_de_Balzac	0.53	0.72	0.81	0.81
dbr:François_Boucher	0.65	0.72	0.81	0.81
dbr:Geer_Van_Velde	0.09	0.09	0.09	0.09
dbr:Ary_Scheffer	0.62	0.72	0.81	0.91
dbr:Jacques_Heim	0.18	0.18	0.45	0.72
dbr:Victor_Hugo	0.53	0.63	0.72	0.91

D. Performances

In this final section, we provide input on the execution performances. Building a model for this work takes 11 minutes on a laptop with 4 GB memory and a quad-core 2.8 GHz processor. We exclude the building of the context graph, which takes place only once. The following time values are only indicative and may vary slightly depending on the parameters chosen to build the model. We use the library rdflib [41] to program the entire process in Python 3. Loading the context graph takes 270 seconds; this only happens once for a set of tests with different parameters to build the model. Generating the 500 walks for each museum takes 30 seconds. Building the model with the walks takes 28 seconds. So, it is possible to generate a model on the fly during interactive sessions.

VII. CONCLUSION

In this paper we presented an approach for link discovery in knowledge bases based on neural language embedding models. We worked on contextualized RDF graphs focused on a specific domain that we extract from large general purpose knowledge graphs. Our approach leverages analogical structures extracted from relational similarities between entities generated by the neural language model. We show how these analogical regularities, applied to entities from knowledge graphs, could be used to infer new unobserved triples from the observed ones. We presented a set of algorithms, on which our approach is based, for the construction of context graphs and the preparation of test data and the evaluation process. The results of applying our approach on a domain-specific ground-truth are promising. We will continue to expand upon the research and compare it with state-of-the-art approaches for knowledge base completion on the standard baselines.

ACKNOWLEDGMENT

This work is conducted as part of the Data & Musée project selected in the $23rd$ call for projects of the Fonds Unique Interministériel (FUI) and certified by Cap Digital and Imaginove. http://datamusee.fr/le-projet/.

REFERENCES

[1] N. Mimouni, J.-C. Moissinac, and A. T. Vu, “Knowledge Base Completion With Analogical Inference on Context Graphs,” in Proceedings of SEMAPRO 2019 : The Thirteenth International Conference on Advances in Semantic Processing, Porto, Portugal, Sep. 2019, pp. 44–47.

[2] M. Al-Ghossein, T. Abdessalem, and A. Barré, “Open data in the hotel industry: leveraging forthcoming events for hotel recommendation,” J. of IT & Tourism, vol. 20, no. 1-4, 2018, pp. 191–216.

[3] S. Szumlanski and F. Gomez, “Automatically acquiring a semantic network of related concepts,” in Proceedings of the 19th ACM CIKM, 2010, pp. 19–28.

[4] J. Yin, X. Jiang, Z. Lu, L. Shang, H. Li, and X. Li, “Neural generative question answering,” in Proceedings of IJCAI. AAAI Press, 2016, pp. 2972–2978.

[5] B. Min, R. Grishman, L. Wan, C. Wang, and D. Gondek, “Distant supervision for relation extraction with an incomplete knowledge base,” in Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Atlanta, Georgia: Association for Computational Linguistics, Jun. 2013, pp. 777–782.

[6] M. Nickel, L. Rosasco, and T. Poggio, “Holographic embeddings of knowledge graphs,” in Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, ser. AAAI’16. AAAI Press, 2016, pp. 1955—-1961.

[7] A. Bordes, N. Usunier, A. García-Durán, J. Weston, and O. Yakhnenko, “Translating embeddings for modeling multi-relational data,” in NIPS, 2013.

[8] H. Liu, Y. Wu, and Y. Yang, “Analogical inference for multi-relational embeddings,” in Proceedings of ICML, 2017, pp. 2168–2178.

[9] A. García-Durán, A. Bordes, N. Usunier, and Y. Grandvalet, “Combining two and three-way embedding models for link prediction in knowledge bases,” J. Artif. Intell. Res., vol. 55, 2016, pp. 715–742.

[10] D. Q. Nguyen, “An overview of embedding models of entities and relationships for knowledge base completion,” CoRR, vol. abs/1703.08098, 2017.

[11] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, “Distributed representations of words and phrases and their compositionality,” in NIPS, 2013.

[12] F. M. Suchanek, G. Kasneci, and G. Weikum, “Yago: A large ontology from wikipedia and wordnet,” Journal of Web Semantics, vol. 6, no. 3, 2008, pp. 203–217, world Wide Web Conference 2007Semantic Web Track.

[13] J. Lehmann, R. Isele, M. Jakob, A. Jentzsch, D. Kontokostas, P. N. Mendes, S. Hellmann, M. Morsey, P. van Kleef, S. Auer, and C. Bizer, “Dbpedia - a large-scale, multilingual knowledge base extracted from wikipedia,” Semantic Web, vol. 6, no. 2, 2015, pp. 167–195.

[14] K. Bollacker, C. Evans, P. Paritosh, T. Sturge, and J. Taylor, “Freebase: A collaboratively created graph database for structuring human knowledge,” in Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data, ser. SIGMOD ’08. New York, NY, USA: Association for Computing Machinery, 2008, pp. 1247—1250.

[15] C. Fellbaum, Ed., WordNet: An Electronic Lexical Database, ser. Language, Speech, and Communication. Cambridge, MA: MIT Press, 1998.

[16] R. Navigli and P. Velardi, “Structural semantic interconnections: a knowledge-based approach to word sense disambiguation,” IEEE transactions on pattern analysis and machine intelligence, vol. 27, no. 7, 2005, pp. 1075–1086.

[17] A. Fader, L. Zettlemoyer, and O. Etzioni, “Open question answering over curated and extracted knowledge bases,” in Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ser. KDD ’14. Association for Computing Machinery, 2014, pp. 1156—-1165.

[18] L. Getoor and B. Taskar, Introduction to Statistical Relational Learning (Adaptive Computation and Machine Learning). The MIT Press, 2007.

[19] Z. Wang, J. Zhang, J. Feng, and Z. Chen, “Knowledge graph embedding by translating on hyperplanes,” in Twenty-Eighth AAAI conference on artificial intelligence, 2014.

[20] Y. Lin, Z. Liu, M. Sun, Y. Liu, and X. Zhu, “Learning entity and relation embeddings for knowledge graph completion,” in Twenty-ninth AAAI conference on artificial intelligence, 2015.

[21] B. Yang, W.-t. Yih, X. He, J. Gao, and L. Deng, “Embedding entities and relations for learning and inference in knowledge bases,” arXiv preprint arXiv:1412.6575, 2014.

[22] T. Trouillon, J. Welbl, S. Riedel, É. Gaussier, and G. Bouchard, “Complex embeddings for simple link prediction.” International Conference on Machine Learning (ICML), 2016.

[23] M. Nickel, V. Tresp, and H.-P. Kriegel, “A three-way model for collective learning on multi-relational data.” in International Conference on Machine Learning (ICML), vol. 11, 2011, pp. 809–816.

[24] R. Socher, D. Chen, C. D. Manning, and A. Ng, “Reasoning with neural tensor networks for knowledge base completion,” in Advances in neural information processing systems, 2013, pp. 926–934.

[25] D. Chen, R. Socher, C. D. Manning, and A. Y. Ng, “Learning new facts from knowledge bases with neural tensor networks and semantic word vectors,” arXiv preprint arXiv:1301.3618, 2013.

[26] H. Liu and Y. Yang, “Cross-graph learning of multi-relational associations,” arXiv preprint arXiv:1605.01832, 2016.

[27] A. Grover and J. Leskovec, “node2vec: Scalable feature learning for networks,” in ACM SIGKDD, 2016.

[28] T. Mikolov, K. Chen, G. Corrado, and J. Dean, “Efficient estimation of word representations in vector space,” in 1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings, Y. Bengio and Y. LeCun, Eds., 2013.

[29] P. Ristoski and H. Paulheim, “Rdf2vec: Rdf graph embeddings for data mining,” in Proceedings of ISWC, 2016, pp. 498 —- 514.

[30] Y. Bengio, R. Ducharme, P. Vincent, and C. Janvin, “A neural probabilistic language model,” J. Mach. Learn. Res., vol. 3, no. null, Mar. 2003, pp. 1137—1155.

[31] Y. Bengio and Y. Lecun, Scaling learning algorithms towards AI. MIT Press, 2007.

[32] T. Mikolov, M. Karafiát, L. Burget, J. Cernocký, and S. Khudanpur, “Recurrent neural network based language model.” in INTERSPEECH, T. Kobayashi, K. Hirose, and S. Nakamura, Eds. ISCA, 2010, pp. 1045–1048.

[33] X. Rong, “word2vec parameter learning explained,” CoRR, vol. abs/1411.2738, 2014. [Online]. Available: http://arxiv.org/abs/1411.2738

[34] D. Meyer, “How exactly does word2vec work?” http://www.1-4-5.net/~dmm/ml/how_does_word2vec_work.pdf, accessed: 2020.05.30.

[35] T. Soru, S. Ruberto, D. Moussallem, A. Valdestilhas, A. Bigerl, E. Marx, and D. Esteves, “Expeditious generation of knowledge graph embeddings,” 2018.

[36] “Rdfs schema,” https://www.w3.org/TR/rdf-schema/, accessed: 2020.05.30.

[37] I. Hulpus, C. Hayes, M. Karnstedt, and D. Greene, “Unsupervised graph-based topic labelling using dbpedia,” in Proceedings of WSDM, 2013, pp. 465–474.

[38] W. Shen, J. Wang, and J. Han, “Entity linking with a knowledge base: Issues, techniques, and solutions,” IEEE Transactions on Knowledge and Data Engineering, vol. 27, no. 2, Feb 2015, pp. 443–460.

[39] “Gensim implementation of word2vec,” https://radimrehurek.com/gensim/models/word2vec.html, accessed: 2020.05.30.

[40] “Paris musées collections website.” http://parismuseescollections.paris.fr/fr/recherche, accessed: 2020.05.30.

[41] “Rdflib python library,” https://github.com/RDFLib/rdflib, accessed: 2020.05.30.

Estimating the Inspection Frequencies of Civil Infrastructures using Correlation Networks and Population Analysis

Prasad Chetti and Hesham H. Ali
College of Information Science and Technology
University of Nebraska at Omaha
Omaha, NE 68182, USA
pchetti@unomaha.edu and hali@unomaha.edu

***Abstract*— Many recent studies have shown that a large percentage of bridges in many parts of the world have a low safety rating. The national bridge inventory database contains data on more than 600,000 bridges, where each bridge has 116 parameters. Current safety inspections require bridge inspectors to manually inspect each bridge every few years. Manpower and budget constraints limit such inspections from being performed more frequently. More efficient approaches need to be developed to improve the process of bridge inspection and increase the overall safety of bridges and civil infrastructures. In this study, we propose a correlation network model to analyze and visualize the big data associated with more than 600,000 bridges in the national bridge inventory database. We use correlation networks based on various safety parameters, then apply the Markov clustering algorithm to analyze a sub-set of 9,546 steel-stringer/multi-beam or girder bridges. We use the produced clusters to propose a different maintenance schedule based on the bridges that show a higher chance of becoming deficient. Results show that of the top five clusters of bridges, three need to be serviced more frequently. We recommend that their inspection frequency be reduced to 12 months instead of 24 months.**

Keywords*— Correlation Networks; Markov Clustering Algorithm; Structural Health Monitoring; National Bridge Inventory database; Inspection Frequency.

I. INTRODUCTION

Every year, the U.S. Federal Highway Administration (FHWA) records the data of more than 600,000 bridges, with a total of up to 116 parameters in the National Bridge Inventory (NBI) database [1][2]. Detailed descriptions of these parameters can be found in the coding guide [3] developed by FHWA. This is a big data and the authors divided some of the parameters associated with these bridges as internal, external, and outcome parameters. For example, the overall fitness rating/safety rating is an outcome rating/parameter of a bridge and is well reflected by the sufficiency rating (SR). The SR ranges between 0 and 1000. The higher the rating, the better the bridge condition is. The Deck Rating (DR), Structural Evaluation Rating (or Structural Condition Rating (SCR)), and Average Daily Traffic (ADT) are some of the internal parameters that affect the outcome rating, such as, the SR. The DR ranges from 0 to 9 and is used to rate the condition of the bridge. The higher the rating, the better the bridge condition is. SCR is calculated based on ADT and other condition ratings and represents the overall structural fitness of the bridge as given in the FHWA coding guide [3]. Ownership (OW) indicates the owner of the bridge responsible for its maintenance, and Inspection Frequency (IF) is an interval usually given in months to indicate how frequently the bridge gets inspected. The latter two are some of the external parameters. Structural Deficiency (SD) is a status assigned to each bridge to indicate whether the bridge is structurally sound or not.

Structural Health Monitoring (SHM) involves implementing a damage detection and characterization strategy for engineering structures [4]. Current safety inspections using traditional SHM mechanisms require bridge inspectors to manually inspect each bridge every few years. Most of the bridges in the NBI database are assigned a safety inspection frequency of 24 months [27]. However, the 24 months' inspection frequency may not be suitable for bridges that require immediate or more frequent attention due to their age or design standards. Manpower and budget constraints limit inspecting the bridges more frequently. Clearly, more efficient approaches need to be developed to improve the process of bridge inspection and increase the overall safety of bridges and civil infrastructures. As a result, we developed a correlation network model (CNM), based on SR rating values of the bridges for 25 years (i.e., from 1992 to 2016) in our earlier conference paper submitted [1]. The main idea behind this work is to use population analysis to assess the health level of each bridge and predict potential health hazards of bridges before they happen. Population analysis means that analyzing different things based on some particular context. Our main hypothesis is that bridges with similar health fitness ratings are included in a common group/cluster in the CNM and have similar behavioral patterns as discussed in our conference paper [1]. As an extension, we further analyzed these individual clusters and assigned updated inspection frequencies based on their structural health and verified what ratings are being affected by ADT.

Our method takes a population of 9,546 steel-stringer/multi-beam or girder design bridges across three states of the USA, such as, California, Iowa, and Nebraska, which come from three different climatic regions [9] of the country as shown in Figure 1. The reason for selecting bridges from different climatic regions is to study whether temperatures play any significant role on the bridges' health in these states.

Initially, we created a correlation network of bridges based on their outcome rating, such as, SR, and then applied the Markov Clustering (MCL) algorithm [10] to produce clusters of bridges. The top 5 clusters were considered for further analysis to see what internal or external parameters enriched to each cluster. In addition, the last 25 years' ADTs were considered to determine their effect on SR and other ratings. The correlations between ADT and other input parameters were also calculated to see what input parameters significantly impacted by ADT in a given cluster. Finally, we visualized the clusters of bridges with respect to their structural deficiency.

As it was already mentioned above, this paper is an extension of our conference paper [1] and here we have conducted and added some other experimental results such as the effect of ADT on various ratings of the bridges, and assigning different inspection frequencies to the bridges in the top 5 clusters. The remainder of this paper is organized as follows. Section II gives the background of this research, where different studies on monitoring bridge health and inspection frequencies are introduced. This section also describes the importance of correlation networks in this study and how correlation networks are used in other domains and ends with a brief introduction of how correlation networks are used in monitoring the structural health of civil infrastructures. Section II also presents some of the key concepts used for creating correlation networks and clusters. Section III describes the methodology used to develop correlation networks and clusters. Section IV presents the experimental results of this study. The final section presents the conclusions and future recommendations.

II. BACKGROUND

Several recent research studies have attempted to estimate the inspection frequencies of highway bridges, and argued that there should be a rationale to set inspection frequencies for both aged bridges and new bridges. While newer bridges may only require inspection 24 months or more, older bridges may need to be inspected more often. Similarly, these studies also argued that inspection frequencies may be rationale when considering different structural condition ratings, design standards and risk factors and proposed an assessment procedure to create inspection intervals for steel bridges with fracture critical members [6]. Some authors have estimated inspection frequencies using life-cycle cost analysis for Stay-Cable replacement design [8]. The Bayesian network model used in [17] demonstrates the predictive and diagnostic capabilities of the model to estimate the load ratings of prestressed concrete bridges and is useful for bridge management. Deterioration models were also developed for Nebraska state bridges using input parameters of the bridges to estimate the deterioration of various condition ratings [5]. Similar studies were done in [24] for developing various deterioration curves using historical data of condition ratings of New York state bridges. A neural network with a novel data organization scheme and voting process used in [28] shows that it can identify damages in bridges with 86% accuracy. Studies also compared various distribution methods to estimating the inspection intervals of bridges using statistical analysis and showed that the Weibull distribution is likely the best fit for historical data of condition ratings [7].

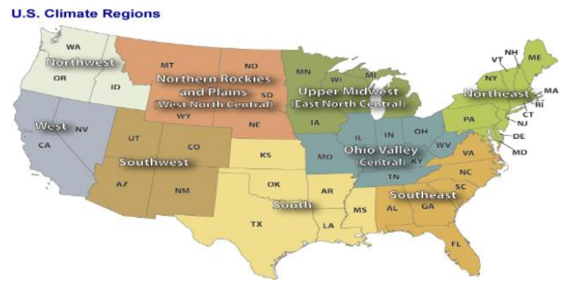


Figure 1. Map of nine USA climate regions (image courtesy : NOAA) [9].

However, all the studies done to date to estimate inspection intervals focused only on individual elements such as deck rating or superstructure rating. Hence, there is a need to estimate inspection intervals based on overall fitness ratings such as SR. If the given set of bridges are clustered based on correlations of historical or time-series data of SR, then individual clusters can be analyzed further to see what parameters are enriched for a cluster of bridges and accordingly the bridge owners may focus on those critical ratings/parameters, and update IF. For example, if a cluster is enriched mostly with structurally deficient bridges, then we can modify the IF of that cluster of bridges to a less than 24 months' interval. Similarly, if a cluster of bridge is enriched with structurally good bridges, then we can update the IF to more than 24 months.

A. Why to utilize a Correlation Network Model?

As the NBI database has data on more than 600,000 bridges, there is a need for powerful and efficient big data tools. CNM is such a powerful big data tool that can predict the structural health of civil infrastructures [1][19]. The key idea behind this work is to use population analysis to assess the health level of each bridge and predict potential health hazards of bridges before they happen. In the population analysis, we compare different clusters or groups of bridges with respect to a particular parameter, based on its enrichment. Our main hypothesis is that the bridges with similar health fitness ratings are included in a common cluster in the CNM and have similar behavioral pattern as shown in [1]. Analyzing these individual clusters will allow us to assign different inspection frequencies based on their deterioration patterns, and structural health.

B. Correlation networks in other disciplines

The ability to show generalization, visualization, and analysis capabilities made the correlation-based network analysis become a powerful analysis tool in biological studies and in various other disciplines [13]. Correlation network analysis was employed in studying biological systems to find the plant growth and biomass in Arabidopsis thaliana Recombinant Inbred Lines (RIL) and introgression lines (IL) [14][15].

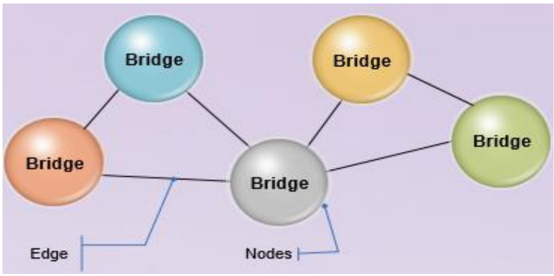


Figure 2. Correlation network representation with bridges.

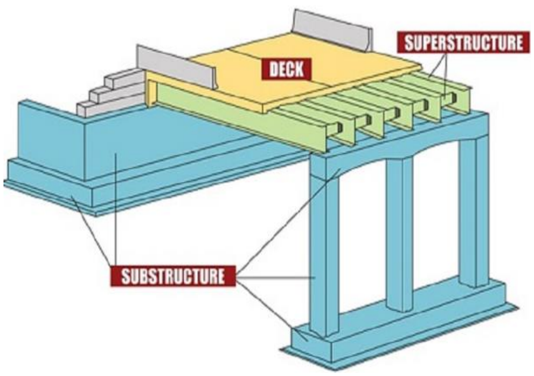


Figure 3. Structural elements of a bridge [22].

It was also successfully applied to evaluate the effect of hypoxia on tumor cell biochemistry [16]. Correlation networks are useful to measure the changes in temporal datasets and study the cluster enrichment by a few Gene Ontology (GO) terms [18].

C. *Correlation Networks for monitoring structural health*

Researchers in the recent past applied CNM to monitor the structural health of civil infrastructures, including analyzing safety issues of various bridges using inventory ratings and other parameters [19]. The basic advantage of using CNM is that the bridges can be clustered together based on some similarity and can be visualized as good and bad bridges [1][19]. As CNM is a new approach for monitoring various civil infrastructures, bridge owners may use CNM to display critical bridges and find an efficient way to improve the inspection schedules [19]. The advantage of this research over [19] is that it considers the temporal-data of SRs. Hence, it can accurately predict an overall fitness rating behavior of the bridges. So, creating a CNM that could deal with temporal-data is one of the objectives of this paper. The motivation of this paper is to process a CNM that could consider bridges' overall behavior (i.e., SR) over a period of time and analyze highly correlated clusters of bridges to predict bridges' behavior and alter the inspection frequencies accordingly. The research question of this paper is to determine what ratings are affected by ADT for each cluster of bridges in the population, if the bridges are clustered using the correlations of temporal data of SRs. The research objective of this paper is to provide a CNM-based decision support System for bridge owners to enable them to find out which bridges need to be serviced first and alter the inspection frequencies. As a result, we developed a novel CNM that considers the temporal data of SRs of the bridges for the last 25 years (from 1992 to 2016), so as to exactly characterize the overall fitness behavior of the bridges over a period of time [1] and see what ratings are effected by ADT, and update the inspection frequencies.

D. *Key concepts used*

The key concepts for this paper are the graph model developing the correlation network, and the Markov clustering algorithm to obtain the group or clusters of bridges.

a) *Graph model*

The CNM is basically an undirected and unweighted graph-based model. The graph is defined as set of vertices and edges, G = (V, E), where V is a set of vertices and E is a set of edges. Each vertex (sometimes called nodes) represents a bridge/civil infrastructure. Two vertices are connected by an edge if and only if their Pearson's correlation coefficient [11] $\rho >= 0.90$, where ρ is a real value, and p-value less than .05. A correlation between any two variables is a value between -1 and +1, which expresses the strength of linkage or co-occurrence. This strength is called Pearson's r or Pearson product-moment correlation coefficient if the correlation is between two continuous–level variables [11][12]. We have used the Pearson's correlation coefficient since the SR data follows normality. Figure 2 represents the undirected and unweighted graph model with 6 nodes and 6 edges. This paper uses bivariate (Pearson's) correlation analysis to show the relationship between any two bridges.

b) *Markov Clustering (MCL)*

In any clustering mechanism, the objects are clustered together in such a way that the distances among clusters are maximized while the distances among the objects are minimized [30] as shown in Figure 4. The Markov Clustering (MCL) algorithm [10] used in this paper is based on the random walks property of the graphs. MCL is a fast and efficient algorithm and is designed for undirected and unweighted graphs. A random walk in a strongly connected cluster usually visits almost all the nodes in the cluster. MCL was already applied on various protein-protein interaction networks and proved to be extraordinarily robust to graph changes and superior in mining complexes from interaction networks [29]. The correlation network that we created for bridges or civil infrastructures is also similar to the protein-protein interaction networks, we employed MCL to mine the groups or clusters of bridges that act similarly.

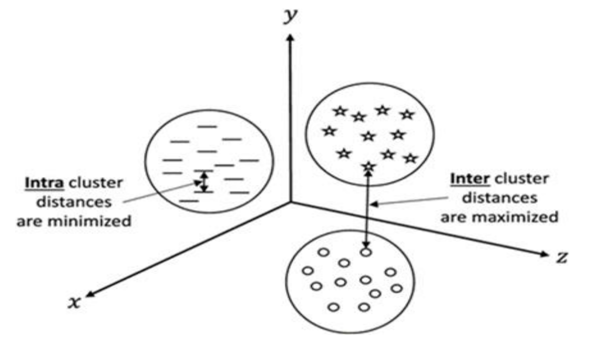


Figure 4. Representation of clustering.

III. METHODOLOGY

The following are the four phases of the CNM that we are proposing.

i. Data acquisition and filtering
ii. Creating a correlation network and applying MCL algorithm
iii. Analyzing various clusters with respect to both input parameters, and output parameters, and comparing various clusters (population analysis)
iv. Developing a decision support system

A. Data Acquisition and Filtering

The highway bridge data is obtained from the NBI database. A total of 25 years (1992-2016) of highway bridge data from three different states of the USA, i.e., California, Iowa, and Nebraska, are considered for this analysis. There are 9,546 highway bridges from these three states, and which are constructed with "Steel material and having Stringer/Multi-beam or Girder design". These are recorded with the numeric 302 in the structure type as given in the FHWA coding guide [3]. The bridges are considered in such a way that their SR is available throughout these 25 years. Hence, each bridge has a minimum age of 25 years. Inconsistent entries, such as the bridges that are recorded as culverts for some years and then non-culverts are removed from the consideration as explained in [1]. Our conference paper explains the data acquisition and filtering in detail [1].

B. Creating a Correlation Network

The SR data of the 9,546 bridges is collected as a matrix along with their 25 years SRs (i.e., from 1992 to 2016). This is called an SR matrix, SR matrix, with each row (i.e., for each bridge) of the matrix having 25 years' SRs in it as a vector. So, there are 9,546 rows in the matrix, A Pearson's correlation coefficients matrix (say, Correlation-matrix) along with the p-values matrix are then obtained over the SR matrix. Each of these matrices are of size 9,546 by 9,546.

Each bridge is then assumed as a node (vertex) in the graph model, and two nodes are connected by an undirected edge if and only if their correlation coefficient $\rho >= 0.90$ and significance value $p < .05$. This creates a correlation network with bridges as nodes along with highly correlated nodes connected by edges as shown in Figure 5. MCL clustering algorithm is then applied in Cytoscape [19] on the previously obtained correlation network to produce clusters. The inflation parameter in MCL clustering can be modified in such a way that the higher inflation value produces clusters of small sizes in terms of nodes. However, we restricted our experiments to the best inflation parameter, such as 1.8, as given in [29]. The clusters produced by MCL algorithm are basically sub-networks of nodes and edges. MCL has produced 8610 nodes in various clusters and 3,865 nodes are present in the top 5 clusters and shown in Figure 6. These top 5 clusters are considered for further analysis. Various experiments are conducted on the top 5 clusters produced by the MCL algorithm, and the results are shown in Section IV below.

C. The parameters considered for analysis

The authors divided the NBI data parameters into three categories, such as, input parameters, output parameters and external parameters—based on their effect on SR as given in the FHWA coding guide [3] to calculate the SR value. SR is an output parameter.

The following are the input and external parameters:

a) *Input parameters:* Average Daily Traffic (ADT), Deck Rating (DR), Superstructure Rating (SpSR), Substructure Rating (SbSR), Structural Condition Rating (SCR), and Water Adequacy Rating (WAR).

b) *External parameters:* State is mentioned as a Location (Loc) (to represent the climatic regions as shown in Figure 1), Owner (OW), Age Category (Age-Cat) derived from the Age (based on Year Built) of the bridge, Inspection Frequency (IF), Rebuilt (RB), and Structural Deficiency (SD) derived from the Status of the bridge.

Some of the above parameters are described as below in the FHWA coding guide [3].

1. Item 1- State Code: considered as the Location (LOC) of the bridge, as the bridges for our analysis were scattered across three states, i.e., California (CA), Iowa (IA), and Nebraska (NE). These three states are from three different climatic regions [9]. California is from the West, Iowa is from the Upper Midwest (East North Central), and Nebraska is from the Northern Rockies and Plains (West North Central) as shown in Figure 1.
2. Item 22- Owner (OW): Maintenance responsibility (Item 21) is used to represent the type of agency

that is the primary owner of the bridge. For example, code '02' in Item 21 is a county highway agency.

3. Item 27- Year Built- records the year that the structure was built. It is used to calculate the age of the bridge. In this study bridges are categorized into three categories based on their age. Ages ranging from 1 to 50 years are in Category A, 51 to 100 years in Category B, and more than 100 years in Category C.
4. Item 29- Average Daily Traffic(ADT): It represents the most recent average daily traffic volume on the bridge.
5. Item 91- Designated Inspection Frequency (IF): It represents the designated inspection frequency of the bridge in months. This interval could be varied from inspection to inspection based on the condition of the bridge. IF=24 indicates that the bridge inspection frequency is for every 24 months as shown in Figure 7.F.
6. Item 106- Year Reconstructed/Rebuilt (RB): It represents the year of reconstruction to keep the bridge operational. RB=0 means that the bridge is not rebuilt. RB=1 indicates that the bridge is rebuilt as shown in Figure 7.E.
7. Status of the bridge: There are four possibilities for the status of the bridges. Status 'N' indicates that it is "Not Applicable". '0' signifies that the bridge is "Not Deficient". '1'indicates that the bridge is "Structurally Deficient (SD)", and '2' means that the bridge is "Functionally Obsolete." A condition rating of 4 or less for Item 58, or 59, or 60, or an appraisal rating of 2 or less for Item 67 or 71, make the bridge structurally deficient [23] as given in [3].
8. Sufficiency rating (SR): It is an outcome measure/rating that is calculated from four factors as given in [3]. It represents the overall fitness rating of the bridge and ranges between 0 and 1000. The lower the rating, the lesser the overall fitness rating is.

IV. EXPERIMENTAL RESULTS

This section demonstrates various experimental results with respect to different network properties, various input, output, and external parameters of the top 5 clusters.

A. Network Properties of Top 5 Clusters

The correlation network (correlation ρ>=0.90) is presented with 9546 nodes, 767542 edges, and 101 connected components. This is basically a scale-free network. In a scale-free network the degree distribution of network follows a power-law. In a power-law node degree distribution, there are many nodes with fewer degrees and fewer number of nodes with more degrees. The nodes with higher degrees could be acting as hub nodes. The study of a hub node is very important with respect to network properties as this hub node is connected to many other similar nodes or bridges. However, studying those hub nodes is beyond the scope of this paper. Figure 6 shows the top 5 clusters (yellow colored clusters) produced by the MCL algorithm. These clusters' statistics are shown in TABLE 1, with the topmost cluster having the highest number of nodes, which is 1,496 and 354,939 edges, and the smallest cluster having 255 nodes and 13,922 edges. The higher the clustering coefficient [23], the higher the degree to which nodes in a graph are inclined to cluster together. The higher values of the average clustering coefficient for each cluster / subnetwork indicate that the nodes inside each cluster tend to be part of that cluster only. Therefore, the top 5 clusters with higher clustering coefficients are considered for further analysis. TABLE 1 shows that cluster 5 has the highest clustering coefficient, which is 0.838. The cluster density describes the potential number of edges present in the sub-network compared to the possible number of edges in the sub-network. From TABLE 1, we see that cluster 3 has the highest density (0.533) among all the top 5 clusters.

B. Population analysis with respect to external parameters

Figures 7.C through 7.F are a comparison of the top 5 clusters with respect to some external parameters such as Age-Category (AGE-CAT), Owner of the bridge (OW), whether the bridge is Rebuilt (RB) or not, and Inspection Frequency (IF). Figure 7.C is a bar chart for comparing the age categories of various clusters. Category A is the set of bridges whose age is 1 to 50 years (labeled blue). B-category bridges are from the age group 51 through 100 (labeled green), and finally the last category, which is Category-C bridges (labeled yellow), with an age of more than 100 years. Cluster 2 is highly enriched with Category-B bridges as shown in Figure 7.C, while the remaining clusters are mostly dominated by both Category-A, and Category B bridges. So, this could have affected the structural deficiency of the bridges as shown in Figure 7.B, where most of the cluster 2 bridges are structurally deficient.

TABLE 1. NETWORK STATISTICS OF TOP 5 CLUSTERS PRODUCED BY THE MCL ALGORITHM.

Cluster Number	#Nodes	#Edges	Avg. Degree	Density	Avg. Clust. Coeff.	SR Avg.
Cluster1	1496	354939	474.51	0.317	0.775	623.7
Cluster2	1180	99000	167.79	0.142	0.674	489.3
Cluster3	634	106955	337.39	0.533	0.823	801.9
Cluster4	300	13377	89.18	0.298	0.812	818.5
Cluster5	255	13922	109.19	0.43	0.838	577.5

Figure 7.D shows that there are six categories of owners of the bridges. However, all the clusters are highly dominated by Owner-2 (labeled green), which is a county highway agency [3]. So, we can infer that all the clusters are enriched and maintained by county highway agency. Figure 7.E is a bar chart that shows whether the bridges were rebuilt (labeled green) or not (labeled blue). If we observe cluster 4, out of 300 bridges, 83 bridges were rebuilt, and most of them were rebuilt in the recent past. The average ratings given in Figures 8.B, and 8.C, show that cluster 4 is a special cluster having lower average ratings in the beginning year 1992 and subsequently increased to high average ratings (especially structural condition rating shown in Figure 8.B) as the number of rebuilt bridges increased in those subsequent years.

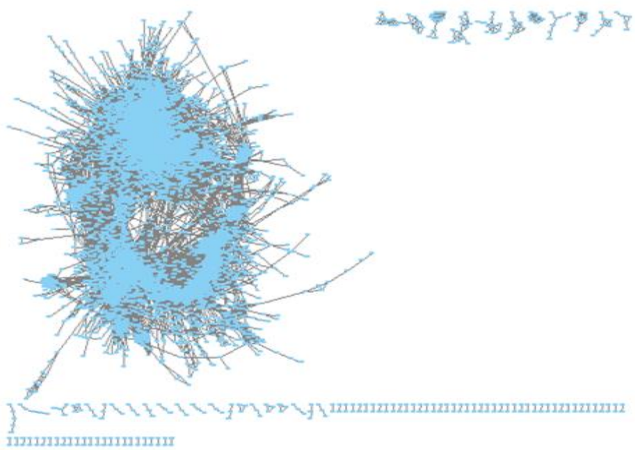


Figure 5. Correlation network (correlation ρ>=*.90*) with 9,546 nodes, and 767,542 edges (Average degree=89.14, and 101 connected components).

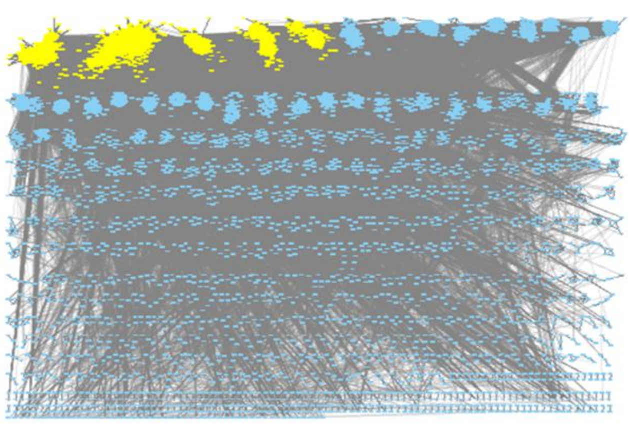


Figure 6. Top 5 clusters (yellow colored clusters) produced by MCL algorithm. (Figure 5 and 6 were generated using Cytoscape [20]).

We also observe from Figures 7.E that most of the cluster 2 bridges are rebuilt, but that have not increased any ratings in Figures 8.B and 8.C. From Figure 7.F, we see that cluster 2 is highly enriched with 24- month inspection frequency. So, we recommend that these bridges' IF's should be lowered, and maintenance must be done more frequently to increase the ratings. An inspection interval of 12 months is more suitable for these bridges.

C. Population analysis of Top 5 clusters with respect to input rating parameters

The input ratings that are considered for population analysis are, DR, SPSR, SBSR, WAR, and SCR of the NBI-dataset-2016. These five different input ratings are compared with respect to their average values in the top 5 clusters as shown in Figure 7. All the average ratings of clusters 3 and 4 are higher compared to all other clusters in top 5. This clearly indicates that these bridges do not have any maintenance issues in terms of any condition ratings in near future. This could be an indication that these bridges' IF's could be updated and increased to either 36 months or 48 months instead of 24 months. Another interesting finding about this figure is that cluster 2 has lower average ratings in all the clusters. For example, cluster 2 has average SBSR value of 4.59. Which is less than 5. According to coding guide [3], these bridges will very soon become structurally deficient. But from Figure 7.F, we see that majority of these bridges' inspection frequencies are 24 months. Hence, we recommend that this cluster's IFs must be lowered to 12 months. Similarly, from Figure 7, if we see the average SCR rating of cluster 2, it is 3.78. which is even below 4. Hence, most of these bridges fell in structurally deficient category. This can be seen from Figure 7.B, as this cluster is highly enriched with SD bridges.

Figures 11 through 15 show the enrichment of individual average input rating values for different clusters. Figure 10 shows that both cluster 1 and 2 are highly enriched with DR = 5. As per the coding guide [3], these bridges are in "Fair Condition" as per the deck is concerned. They are just 1 rating above structural deficiency rating, such as rating 4. This calls the frequent maintenance, such as making IF=12 months for the decks of the bridges in these clusters.

Figures 12 and 13 show the average condition ratings of superstructures and substructures, respectively. From Figure 11, we see that cluster 2 is enriched with SpSR <= 5. Once these bridges' SpSR's drop from 5 to 4, then most the bridges will fall into the SD bridge category. Hence, the improvement in the SpSR rating in terms of dropping the live load is required. This can be done by reducing Average Daily Traffic and implementing required recovery services on these bridges. Cluster 2 from Figure 12 also shows that the substructure rating (SBSR) is critical, as most of the bridges' enriched with SBSR ratings <= 5. Figure 13 shows that all the clusters are highly enriched with water adequacy rating (WAR) >=6. Hence, any improvement in terms of WAR rating is not required for the next couple of years.

From Figure 14, we find that the bridges in clusters 1, 2, and 5 are highly enriched with average structural condition rating (SCR) <=4, which shows that the maximum of the bridges in these clusters are either structurally deficient (SD) or very soon will they fall into this SD category. Hence, more frequent (such as, IF=12 months) maintenance is required for these clusters.

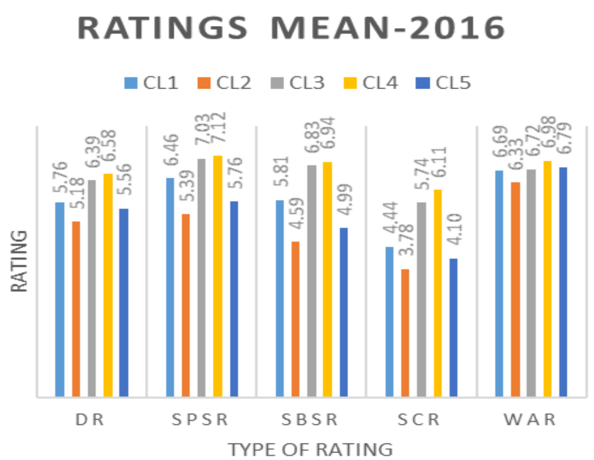


Figure 7. Ratings comparison for top 5 clusters (year 2016).

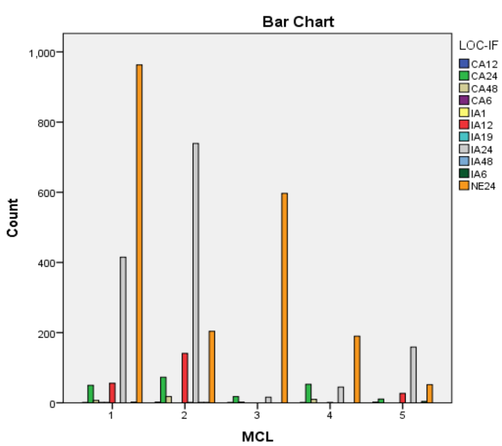


Figure 7.A. Location-Inspection Frequency (Loc-IF).

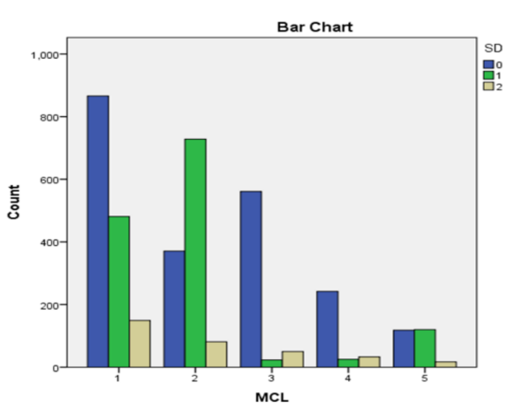


Figure 7.B. Structural Deficiency (SD).

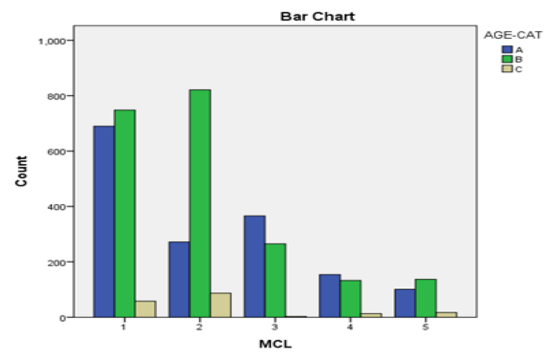


Figure 7.C. Age-Category (AGE-CAT).

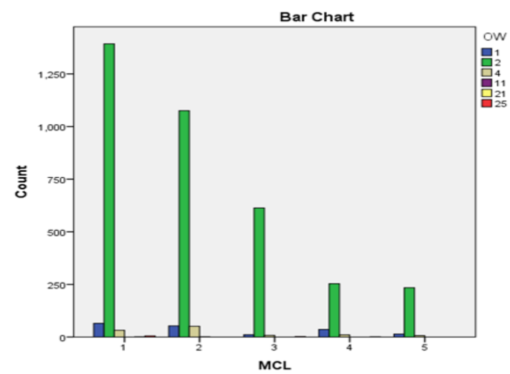


Figure 7.D. Owner(OW).

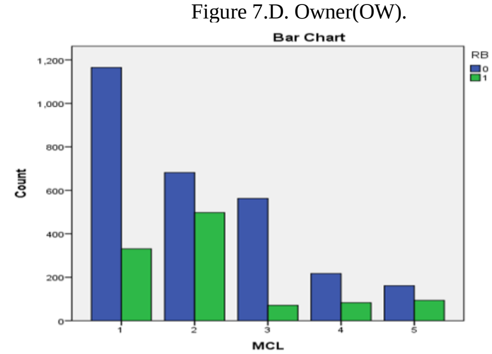


Figure 7.E. Rebuilt (RB).

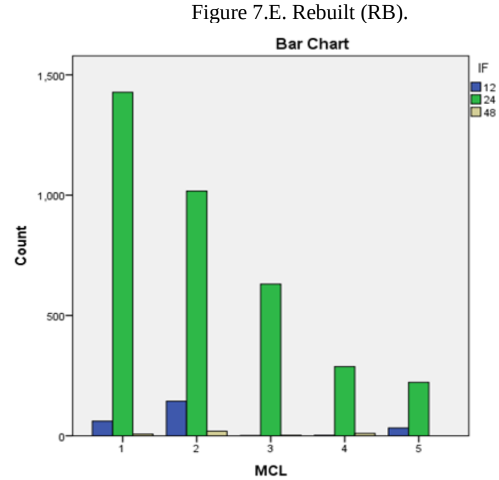


Figure 7.F. Top 5 clusters' Inspection Frequencies (IF). (All these bar charts from Figures 7.A through 7.F were produced in SPSS [25]).

D. Effect of the Average Daily Traffic(ADT) on various ratings

Figures 8.A, through 8.C show the averages of Averages Daily Traffic (ADT), Structural Condition Rating (SCR), and Sufficiency rating (SR), respectively, for the years 1992, 1997, 2002, 2007, 2012, and 2016 (i.e., for every 5 years except for the year 2016, which has a gap of four years from 2012). The ADT graph shown in Figure 8.A indicates that the traffic rate is increasing year by year for any cluster of bridges except for the cluster 2 (between the years 1997 and 2002) and for the cluster 5 (between the years 2007 and 2012). Which clearly indicates that some measures regarding lowering the ADT must have been taken in those clusters in those years to increase the life of the bridges. However, that is not enough to increase the ratings of these bridges at the end year 2016. But at the same time, we see that the increase in ADT, decreases the ratings in case of remaining clusters as we compare Figures 8.A, with 8.B, and 8.C. As these input ratings are reduced, it automatically affects the SR (the overall fitness) and hence, it gets reduced. For clusters 1, 2 and 5, the effect of ADT is huge on both input ratings (for example, SCR), and on output rating SR. As most of the bridges in these three clusters are already structurally deficient, the effect of ADT would be instant. From Figure 8.D, we see that mean ADT volumes increased only due to category D (ADT>=5000) in cluster 1. This indicates that the very high traffic needs to be controlled for the bridges in cluster 1.

E. Analysis with respect to Structural Deficiency

The Structural Deficiency (SD) of the top 5 clusters is shown in Figure 7.B and their individual visualized comparison is shown in Figures 9.A through 9.E, with a total of 61.69% of cluster 2, 47.06% of cluster 5, and 32.15% of cluster 1 bridges are structurally deficient. At the same time, further analysis on cluster 2, as shown in Figures 9.B and 9.F, shows that this cluster is enriched with 62.63% Iowa and 17.29% Nebraska bridges, and both have 24-month inspection frequencies and a mean sufficiency rating below 500. The same is the case for cluster 1. Here, more than 60% of the bridges are from Nebraska with 24-month inspection frequency and having a mean sufficiency rating of just above 600. Hence, we suggest that the 24-month inspection frequency for these structurally deficient bridges needs to be modified to a 12-month inspection frequency to provide rehabilitation services more frequently. Similarly, only 3.63% and 8.33% of cluster 3 and cluster 4 bridges are structurally deficient, as shown in Figures 9.C and 9.D. Cluster 3 and Cluster 4 are enriched with more than 94% and 63% of Nebraska bridges, respectively, with 24-month inspection frequency (as shown in Figure 7.A). Also, the average sufficiency rating of these two clusters is above 800 as shown in Figure 15. Hence, these bridges' 24-month IF can be increased to either 36 months or 48 months as these bridges need not be serviced more frequently.

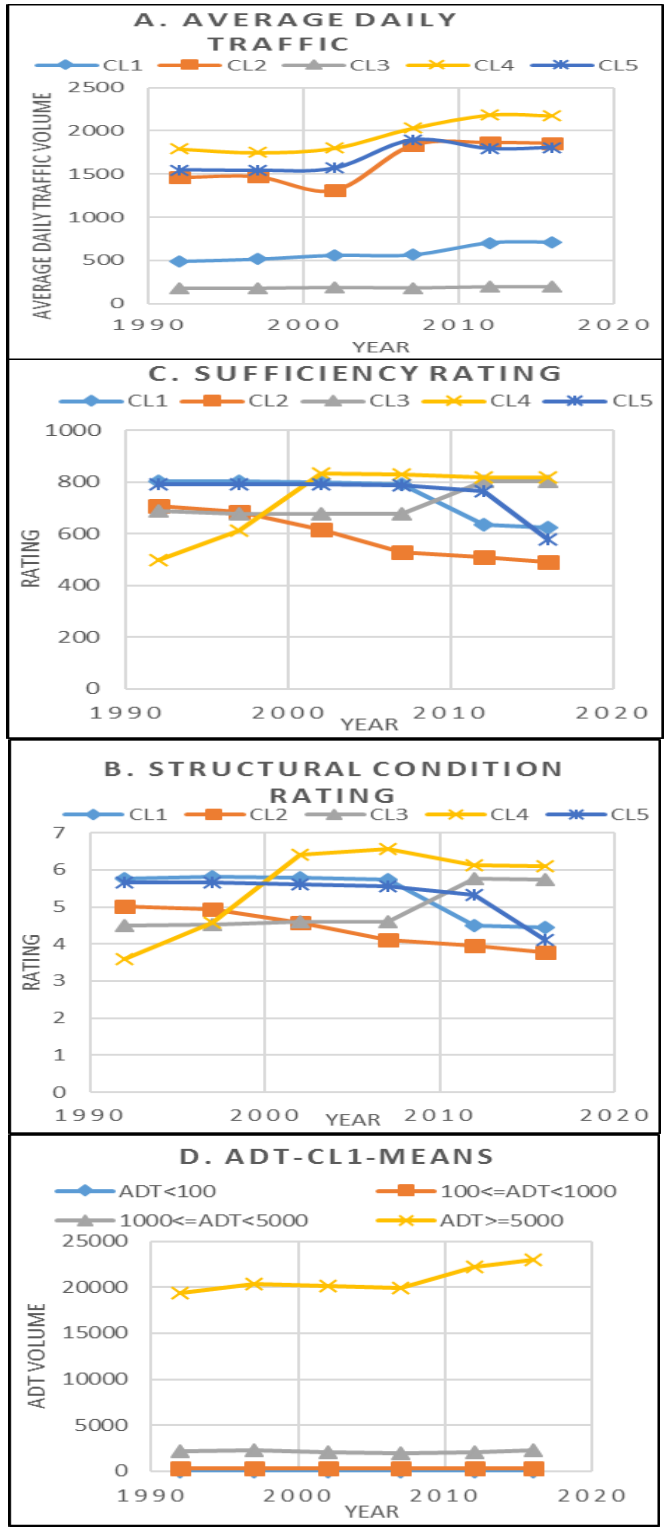


Figure 8. Comparison of top 5 clusters' averages (dataset years 1992,1997,2002,2007,2012 and 2016) with respect to various ratings and average daily traffic. A: Average Daily Traffic (ADT). B: Structural Condition Rating (SCR). C: Sufficiency rating (SR). D: Cluster 1 means of ADT, category wise (four different colors indicating four different categories. Category-A: ADT<100, B: 100<=ADT<1000, C: 1000<=ADT<5000, and category D: ADT>=5000).

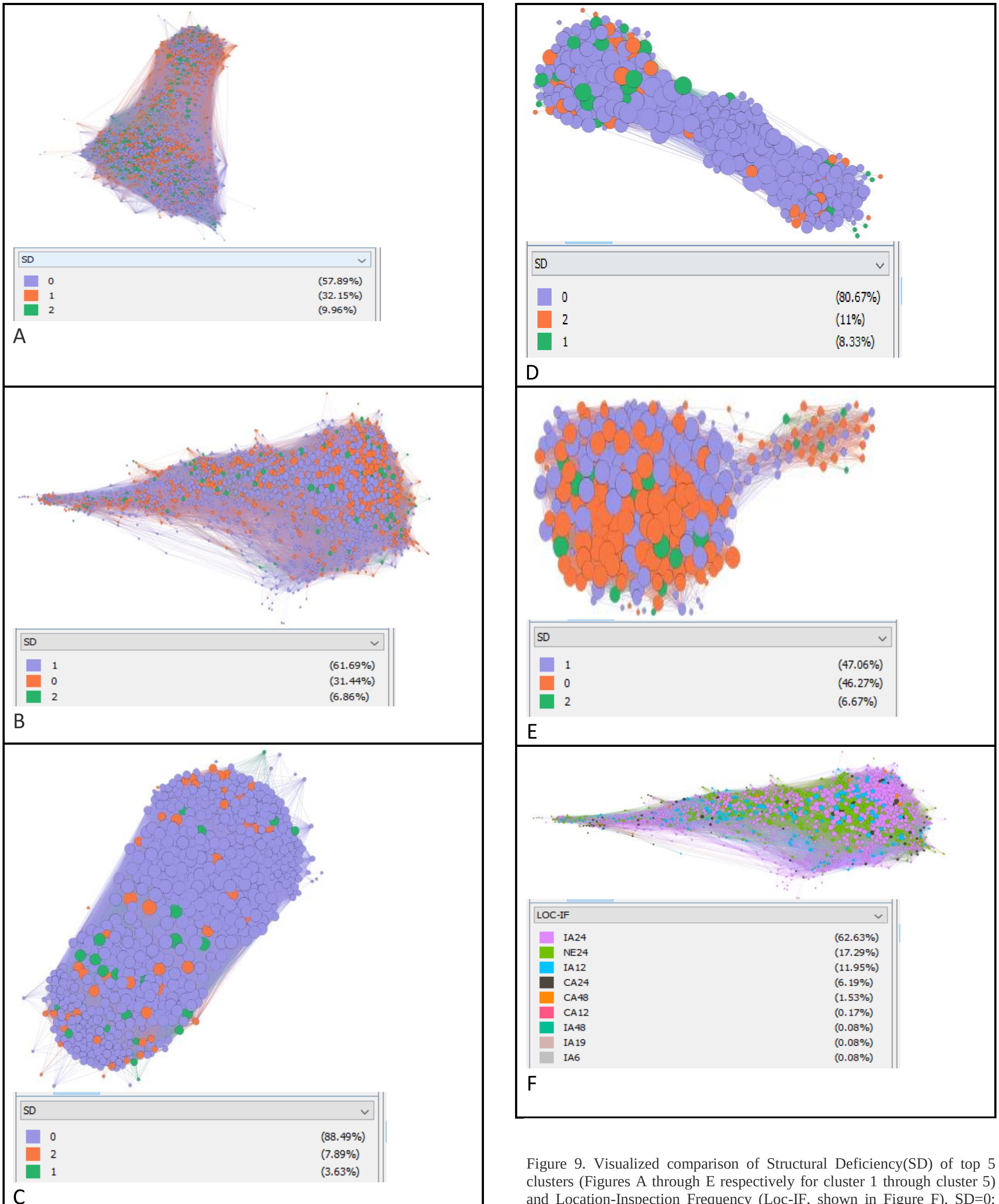


Figure 9. Visualized comparison of Structural Deficiency(SD) of top 5 clusters (Figures A through E respectively for cluster 1 through cluster 5) and Location-Inspection Frequency (Loc-IF, shown in Figure F). SD=0: Structurally Good, SD=1: Structurally Deficient, SD=2: Functionally Obsolete. (The above images were generated using Gephi [21]).

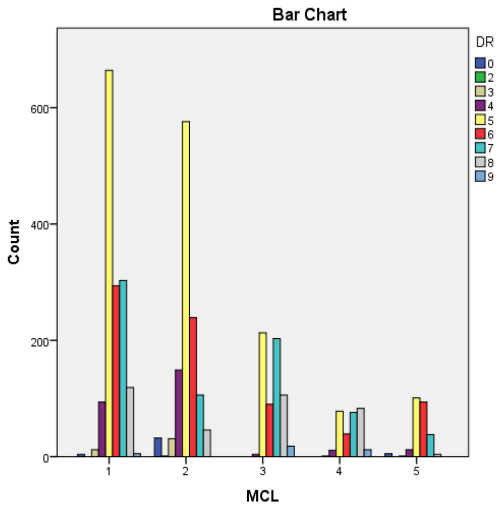


Figure 10. Comparison of top 5 clusters with respect to Deck Rating (DR).

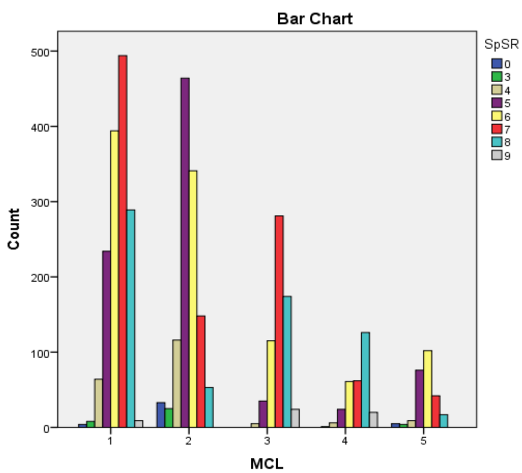


Figure 11. Comparison of top 5 clusters with respect to Superstructure Rating (SpSR).

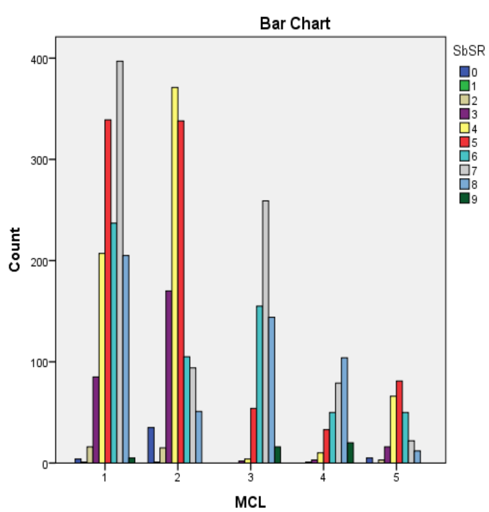


Figure 12. Comparison of top 5 clusters with respect to Substructure Rating (SBSR).

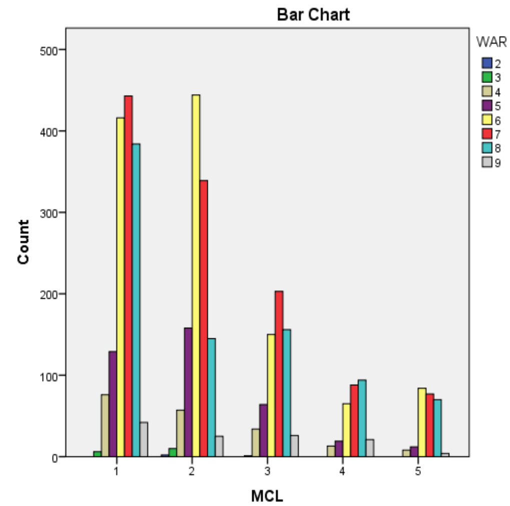


Figure 13. Comparison of top 5 clusters with respect to Water Adequacy Rating (WAR).

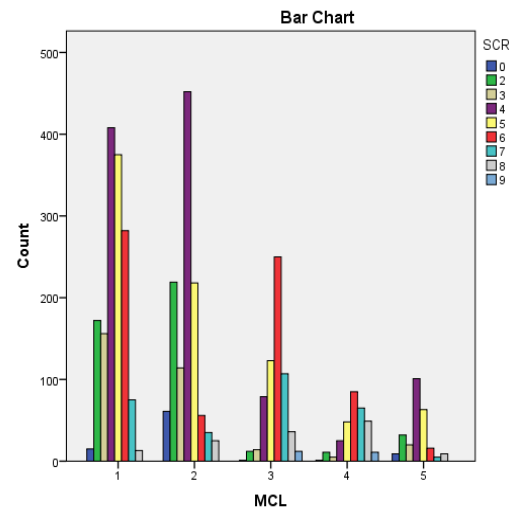


Figure 14. Comparison of top 5 clusters with respect to Structural Condition Rating (SCR).

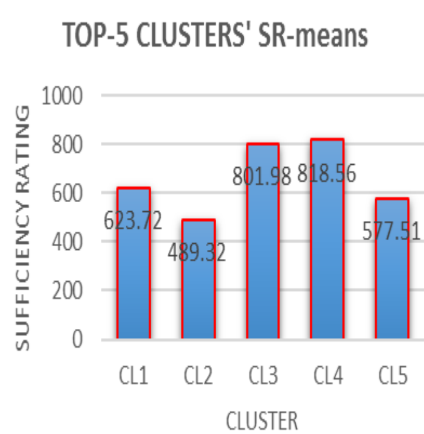


Figure 15. Averages of SRs of top 5 clusters for the year 2016.

V. CONCLUSION

We have presented a model for identifying which bridges need to be serviced first and suggested intervals of inspection using correlation networks and population analysis along with cluster enrichment parameters. Different clusters are enriched with different parameters and/or ratings. Each cluster is analyzed and has given sufficient evidence for our original hypothesis that bridges with similar characteristics/parameters must be part of the same cluster. We found that there are three clusters of bridges (clusters 1,2 and 5), which have low values of sufficiency ratings as shown in Figure 15, and to which immediate rehabilitation services are required. Hence, they need to be serviced first. Therefore, their inspection frequency should be adjusted from 24 months to 12 months. Out of these three clusters, cluster 2 is eligible for federal funding as the average SR for this cluster is below 500 (SR value below 500 makes a bridge eligible for federal funding as per FHWA guidelines [2]) as shown in Figure 15. From Figure 15, we also found that the bridges in clusters 3 and 4 have high overall fitness ratings. Therefore, they do not require immediate attention and their inspection frequencies could be adjusted to 36 months or 48 months instead of 24 months.

We have further analyzed the clusters to see how increased ADT leads to decreased overall fitness (in terms of SR) of the bridges. We have also presented visualizations of all the top 5 clusters with respect to SD to allow different bridge owners to clearly distinguish what bridges are deficient, or functionally obsolete, or in good condition.

From various visualizations and statistics with respect to different input ratings and external parameters, our decision support system could visualize that most of the bridges (built with steel material and having stringer/multi-beam or girder design) from clusters 1, 2 and 5 are in both Iowa and Nebraska states, and with IF = 24 months, and aged above 50 years. These bridges need to be serviced first and their inspection frequencies need to be adjusted to 12 months instead of 24 months.

With all these results, our correlation network model enables various bridge authorities to clearly distinguish between the structurally good and deficient bridges. SHM inspectors can now estimate which bridges' IFs need to be adjusted to 12 months instead of 24 months. Rehabilitation services should be provided accordingly, and authorities can distribute funds on priority basis which could result in saving money and many human lives.

One shortfall of this method is that the big data associated with the information of thousands of bridges may consume more time to create the correlation matrix and correlation network, but with the power of existing supercomputers and their huge memories this couldn't be a big problem.

Presently we have considered only the bridges that are constructed with steel material and with stringer / multi-beam or girder design. As a future work, we may study CNM to see the clustering of various bridges constructed with a different design and/or with different material, and for different states. We would also like to study these clusters further and assign risk rankings to these top 5 clusters, to prioritize the clusters that need immediate rehabilitation services. Further studies may also focus on studying the remaining clusters instead of top 5 clusters. One can also use either DR or SbSR instead of SR, to create the correlation network and provide the inspection frequencies based on the temporal data of these ratings. We can also study the network properties of these clusters in detail to get more insights about these groups of bridges. This work may further be extended for verifying the temperatures role on the bridges as they come from different climatic regions.

ACKNOWLEDGEMENTS

The authors of this paper would like to thank the administration of the University of Nebraska at Omaha, USA as this research work is sponsored by the university under the GRACA research grant with the application number 16796.

REFERENCES

[1] P. Chetti and H. H. Ali. "Analyzing the Structural Health of Civil Infrastructures Using Correlation networks and Population Analysis", Proceedings of the Eighth International Conference on Data Analytics, Porto, Portugal, 2019.

[2] Federal Highway Administration (FHWA), https://www.fhwa.dot.gov/bridge/nbi/ascii.cfm , 2018.08.15

[3] Federal Highway Administration (FHWA), https://www.fhwa.dot.gov/bridge/mtguide.pdf, 2018.08.15

[4] C. R. Farrar and K. Worden, "An introduction to structural health monitoring", Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 365.1851, pp. 303-315, 2007.

[5] A. Hatami and G. Morcous, "Developing deterioration models for Nebraska bridges", No. M302, 2011.

[6] J. Parr, Michael, J. Conner, Robert, and Mark Bowman. "Proposed method for determining the interval for hands-on inspection of steel bridges with fracture critical members", *Journal of Bridge Engineering*, 15.4 , 2009, pp. 352-363.

[7] M. Nasrollahi and G. Washer, "Estimating inspection intervals for bridges based on statistical analysis of national bridge inventory data," Journal of Bridge Engineering, Vol. 20, No. 9, 04014104, 2014.

[8] A. B. Mehrabi et al., "Evaluation, rehabilitation planning, and stay-cable replacement design for the hale boggs bridge in Luling, Louisiana", *Journal of Bridge Engineering*, 15.4, 2010, pp. 364-372.

[9] T. Karl and W.J. Koss, "Regional and national monthly, seasonal, and annual temperature weighted by area, 1895-1983", 1984.

[10] S. Dongen, "A cluster algorithm for graphs", CWI (Centre for Mathematics and Computer Science), 2000.

[11] K. Pearson, "On the coefficient of racial likeness", Biometrika pp. 105-117, 1926.

[12] J. Benesty et al., "Pearson correlation coefficient", Noise reduction in speech processing, Springer Berlin Heidelberg, pp. 1-4, 2009.

[13] A. Batushansky, D. Toubiana and A. Fait, “Correlation-based network generation, visualization, and analysis as a powerful tool in biological studies: a case study in cancer cell metabolism”, BioMed research international, 2016.

[14] J. Lisec et al., “Identification of metabolic and biomass QTL in Arabidopsis thaliana in a parallel analysis of RIL and IL populations”, The Plant Journal, Vol. 53, No. 6, pp. 960-972, 2008.

[15] R.C. Meyer et al., “The metabolic signature related to high plant growth rate in Arabidopsis thaliana”, Proceedings of the National Academy of Sciences, Vol. 104, No. 11, pp. 4759-4764, 2007.

[16] H.L. Kotze, E.G. Armitage, K.J. Sharkey, J.W. Allwood, W.B. Dunn, K.J. Williams and R. Goodacre, “A novel untargeted metabolomics correlation-based network analysis incorporating human metabolic reconstructions”, BMC systems biology, Vol. 7, No. 1, pp. 107, 2013.

[17] K. LeBeau and W.F. Sara, "Predictive and diagnostic load rating model of a prestressed concrete bridge", *Journal of Bridge Engineering*, 15.4, 2010, pp. 399-407.

[18] K. Dempsey, I. Thapa, D. Bastola and H. Ali, “Identifying modular function via edge annotation in gene correlation networks using Gene Ontology search”, Bioinformatics and Biomedicine Workshops (BIBMW), IEEE International Conference, pp. 255-261, 2011.

[19] A. Fuchsberger and H. Ali, "A Correlation Network Model for Structural Health Monitoring and Analyzing Safety Issues in Civil Infrastructures", Proceedings of the 50th Hawaii International Conference on System Sciences, 2017.

[20] P. Shannon et al., “Cytoscape: a software environment for integrated models of biomolecular interaction networks”, Genome research, Vol. 13, No. 11, pp. 2498-2504, 2003.

[21] M. Bastian, S. Heymann, and M. Jacomy, “Gephi: an open source software for exploring and manipulating networks”, ICWSM, Vol. 8, pp. 361-362, 2009 (retrieved on August 15, 2017 from https://gephi.org/)

[22] Daily Civil, http://www.dailycivil.com/structural-elements-bridge/ 2018.08.15

[23] Federal Highway Administration (FHWA), https://www.fhwa.dot.gov/bridge/britab.cfm 2018.08.15

[24] A.K. Agrawal, K. Akira, C. Zheng. "Deterioration rates of typical bridge elements in New York." *Journal of Bridge Engineering* 15.4 (2010): 419-429.

[25] A. Field, “Discovering statistics using IBM SPSS statistics”, Sage, 2013.

[26] http://journal.r-project.org/archive/2013-1/kahle-wickham.pdf

[27] C.F.R., *National Bridge Inspection Standards*, in *23 CFR part 650*. 2004: USA. p.7441974439.

[28] Li Zhe and R. Burgueño, "Using soft computing to analyze inspection results for bridge evaluation and management", *Journal of Bridge Engineering*, 15.4, 2010, pp. 430-438.

[29] S. Brohee and J. V. Helden. "Evaluation of clustering algorithms for protein-protein interaction networks." BMC bioinformatics, Vol. 7, No. 1, pp. 488, 2006.

[30] A.K. Jain, “Data clustering: 50 years beyond K-means”, Pattern recognition letters, Vol. 31, No. 8, pp. 651-666, 2010.