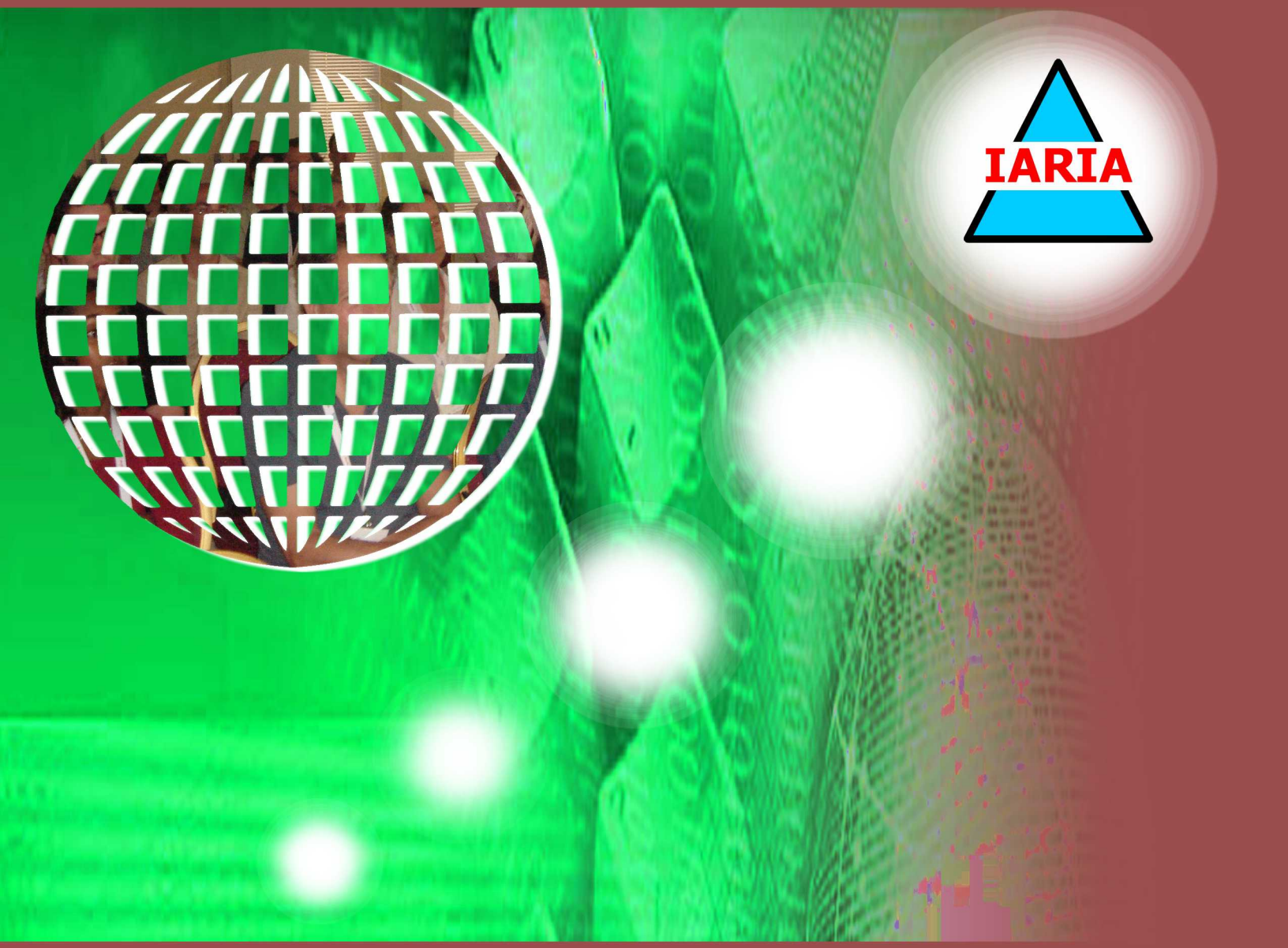
International Journal on Advances in Life Sciences



2020 vol. 12 nr. 1&2

The *International Journal on Advances in Life Sciences* is published by IARIA.
ISSN: 1942-2660
journals site: http://www.iariajournals.org
contact: petre@iaria.org

Sponsored by IARIA
www.iaria.org

International Journal on Advances in Life Sciences
Volume 12, Number 1 & 2, 2020

Dickson Chiu, Dickson Computer Systems, Hong Kong
Bee Bee Chua, University of Technology, Sydney, Australia
Udi Davidovich, Amsterdam Health Service - GGD Amsterdam, The Netherlands
Maria do Carmo Barros de Melo, Telehealth Center, School of Medicine - Universidade Federal de Minas Gerais (Federal University of Minas Gerais), Brazil
Kari Dyb, Norwegian Centre for Integrated Care and Telemedicine / University Hospital of North Norway | University of Tromsø, Norway
Juergen Eils, DKFZ, German
Anne G. Ekeland, Norwegian Centre for Integrated Care and Telemedicine / University Hospital of North Norway | University of Tromsø, Norway
El-Sayed M. El-Horbaty, Ain Shams University, Egypt
Ivan Evgeniev, TU Sofia, Bulgaria
Karla Felix Navarro, University of Technology, Sydney, Australia
Joseph Finkelstein, The Johns Hopkins Medical Institutions, USA
Stanley M. Finkelstein, University of Minnesota - Minneapolis, USA
Adam M. Gadomski, Università degli Studi di Roma La Sapienza, Italy
Ivan Ganchev, University of Limerick, Ireland / University of Plovdiv “Paisii Hilendarski”, Bulgaria
Jerekias Gandure, University of Botswana, Botswana
Xiaohong Wang Gao, Middlesex University - London, UK
Josean Garrués-Irurzun, University of Granada, Spain
Hassan Ghazal, Moroccan Society for Telemedicine and eHealth, Morocco
Piero Giacomelli, SPAC SPA -Arzignano (Vicenza), Italia
Alejandro Giorgetti, University of Verona, Italy
Anthony Glascock, Drexel University, USA
Wojciech Glinkowski, Polish Telemedicine Society / Center of Excellence "TeleOrto", Poland
Francisco J. Grajales III, eHealth Strategy Office / University of British Columbia, Canada
Conceição Granja, Conceição Granja, University Hospital of North Norway / Norwegian Centre for Integrated Care and Telemedicine, Norway
William I. Grosky, University of Michigan-Dearborn, USA
Richard Gunstone, Bournemouth University, UK
Amir Hajjam-El-Hassani, University of Technology of Belfort-Montbéliard, France
Lynne Hall, University of Sunderland, UK
Päivi Hämäläinen, National Institute for Health and Welfare, Finland
Anja Henner, Oulu University of Applied Sciences, Finland
Marike Hettinga, Windesheim University of Applied Sciences, Netherlands
Stefan Hey, Karlsruhe Institute of Technology (KIT) , Germany
Dragan Ivetic, University of Novi Sad, Serbia
Sundaresan Jayaraman, Georgia Institute of Technology - Atlanta, USA
Malina Jordanova, Space Research & Technology Institute, Bulgarian Academy of Sciences, Bulgaria
Attila Kertesz-Farkas, University of Washington, USA
Hassan Khachfe, Lebanese International University, Lebanon
Valentinas Klevas, Kaunas University of Technology / Lithuaniain Energy Institute, Lithuania
Anant R Koppar, PET Research Center / KTwo technology Solutions, India
Bernd Krämer, FernUniversität in Hagen, Germany
Ramesh Krishnamurthy, Health Systems and Innovation Cluster, World Health Organization - Geneva, Switzerland
Roger Mailler, University of Tulsa, USA
Dirk Malzahn, OrgaTech GmbH / Hamburg Open University, Germany
Salah H. Mandil, eStrategies & eHealth for WHO and ITU - Geneva, Switzerland
Herwig Mannaert, University of Antwerp, Belgium
Agostino Marengo, University of Bari, Italy
Igor V. Maslov, EvoCo, Inc., Japan
Ali Masoudi-Nejad, University of Tehran , Iran
Cezary Mazurek, Poznan Supercomputing and Networking Center, Poland

International Journal on Advances in Life Sciences
Volume 12, Numbers 1 & 2, 2020

CONTENTS

A Monitoring System for Operating Theaters at Heidelberg University Hospital

First Experiences Implementing Predictive Analytics Tools in a Clinical Routine Setting

Oliver Klar
Department of Medical Information Systems
University Hospital Heidelberg
Heidelberg, Germany
oliver.klar@med.uni-heidelberg.de

Gerd Schneider
Department of Medical Information Systems
University Hospital Heidelberg
Heidelberg, Germany
gerd.schneider@med.uni-heidelberg.de

Rasim Atakan Poyraz
Department of Medical Information Systems
University Hospital Heidelberg
Heidelberg, Germany
atakan.poyraz@med.uni-heidelberg.de

Oliver Heinze
Department of Medical Information Systems
University Hospital Heidelberg
Heidelberg, Germany
oliver.heinze@med.uni-heidelberg.de

Abstract - **The rise of Artificial Intelligence (AI) is ubiquitous. In healthcare it is seen as a key technology supporting clinicians in their daily routine. The PART research project (Predictive Analytics of Robustness Testing) aims to develop an AI driven, vendor independent monitoring system, which has the focus on system monitoring, profitability analysis, and predictive maintenance of networked medical devices in a clinical environment. However, before working on AI driven monitoring solutions at Heidelberg University Hospital, we experienced a variety of difficulties according to networked medical devices, data acquisition, standards and protocols, and device interfaces, which must be addressed first. This paper stresses those difficulties and presents a monitoring system of networked medical devices from one operating theater at Heidelberg University Hospital. Continuous data streams of laparoscopic devices out of the surgery room are ingested into the system and analyzed in real-time. The results are stored in an on-premises data store and visualized according to profitability analysis and system monitoring in a dashboard. Further, an outlook is giving including the transformation of the presented monitoring system into the Medical Data Integration Center (MeDIC) of the Heidelberg University Hospital in the future and the connection of more surgery theaters.**

Keywords - clinical artificial intelligence; artificial intelligence in healthcare; medical device monitoring; real-time data stream processing; predictive maintenance; Apache Kafka; Apache Flink; Elasticsearch; Kibana.

I. BACKGROUND

The PART research project (Predictive Analytics of Robustness Testing) aims to develop an AI driven, vendor independent monitoring system, which has the focus on system monitoring, profitability analysis, and predictive maintenance of networked medical devices. However, at Heidelberg University Hospital we experienced a variety of difficulties building such a monitoring system including data acquisition, standards and protocols and device interfaces [1]. Dealing with those circumstances in this work, we present a flexible and extendable pipeline architecture for ingestion, processing, and storage of medical device data. As a first approach we focus on monitoring laparoscopic medical devices of our project partner Karl Storz GmbH & Co. KG from one operation theater. Use cases for this system from the perspective of the Heidelberg University Hospital are:

- Profitability Analysis

Acquisition and operation of a vast number of medical devices is expensive. Often several devices of the same type are used in the same clinic. So far, there are no numbers about device usage and whether the current number of devices is required. The monitoring system should collect key figures such as utilization and operating hours, which are then economically analyzed.

- Online Inspection

To conduct inspections on medical devices, such as safety related checks and metrological checks, it is necessary to put them out of operation. Those checks must be done in fixed predefined intervals and are known as preventive maintenance [2]. To reduce this costly downtime, the goal of the monitoring system is to go from preventive maintenance to predictive maintenance by using predictive analytics tools to determine when maintenance actions are required automatically.

Analyzing vast amounts of medical device data continuously, machine learning (ML) algorithms should

help putting medical devices out of order only when thresholds are overstepped, or critical errors occur. This should help avoiding unnecessary down-times, optimizing the maintenance schedules, and reducing maintenance costs.

- System Monitoring

To exchange data with clinical systems like a Picture Archiving and Communication System (PACS) [3] or Hospital Information Systems (HIS), medical devices are interconnected more often. To guarantee stability for such a growing networked system and manage a reliable exchange of data between devices and other IT infrastructures, a superior overall system is required for monitoring and alerting.

This work presents first experiences in designing and implementing such a vendor-independent monitoring system, facing the real world setting of a university hospital.

In Section II we describe the challenges we face implementing such a monitoring system. The PART pipeline architecture of our monitoring system is the focus of Section III. The dashboard of the monitoring system is presented in Section IV and we conclude with discussion and outlook in Section V.

II. Challenges & Objectives

There are several hurdles according to healthcare devices, medical device data, data processing, and data protection that must be taken in advance of realizing a powerful monitoring system. First, one must address problems caused by heterogeneity. Devices are mostly, due to reasons of independence, from different manufacturers. This ranges from infusion pumps to the latest CT or MRI scanners. Even though there are standards for networked medical devices in operation rooms (e.g., IEEE 11073), they are only implemented and promoted by a few manufacturers [4]. The communication of the most medical equipment works mostly via proprietary interfaces and protocols, and manufacturers are very reluctant disclosing those interfaces, or implementing given standards. Even if monitoring of those medical devices is possible in general, integrated sensors like in the Philips e-Alert System of MRI scanners [5] or in industrial environments, which produce data feasible according to predictive maintenance, are rather seldom and hence can restrict available information to log data of the machines [6]. Further, expensive devices like CT scanners usually have extensive maintenance contracts, which include that maintenance, repair, and service may only be performed by a service engineer of the manufacturer itself. Collecting relevant data from such devices, e.g., getting information about the condition and operating status is demanding.

Those manufacturers come up with own solutions for monitoring the medical device fleet [7][8]. They provide the customer with key performance indicators which help e.g., identifying over- and under-utilized devices, balancing the product use and lessen the strain placed on individual medical devices [9].

As the operator of all medical equipment, Heidelberg University Hospital ends up maintaining a monitoring system for each manufacturer, which is not expedient.

Another potential source of data, delivering information about the status of medical equipment, could be the usage of IoT sensors. However, gathering data by additional attached sensors in a sterile environment like an operating room is under serve restrictions due to aspects like patient safety. Hence, the situation is barley comparable to an industrial production line where predictive maintenance is quite common for prevention device failures.

Data quality is a problem since several decades. In contrast to big data, machine learning goes through with a different set of data quality concerns. The three components of ML algorithms are model representation, measures of valuating model accuracy, and methods for searching the best ML model [10]. Since these three components are highly related to each other, data quality for ML is very complex. One of the biggest concerns in big data is missing data as well as the well-structured datasets.

In PART the current question is not which data mining algorithms fit the most for our needs, the question is where the data is coming from in the first place. Therefore, we are looking in all directions and started working with simulated device data as well as getting familiar with the data mining approaches. Another subject is data protection and privacy. By monitoring devices, collecting, and analyzing data, it could be possible to draw conclusions about patients, treatment, and the work of clinical personnel itself. This is sometimes seen very critically by the clinic staff and requires a close examination and further steps like anonymization of the data.

Finally, an important point to mention is that according to predictive maintenance of networked medical devices, failures are quite often due to simple reasons like dropping the device or too much moisture when cleaning in a clinical environment. Two issues which are hard to handle by analyzing device data.

Although devices in clinical environments produce a high volume of data, it is quite challenging, as described above, to access, evaluate, and generate added value from this data treasure.

Keeping those challenges in mind in our first approach we have focused on descriptive analytics towards the development of an AI driven monitoring tool including ingestion, real-time analytics, storage, and visualization of networked medical device data.

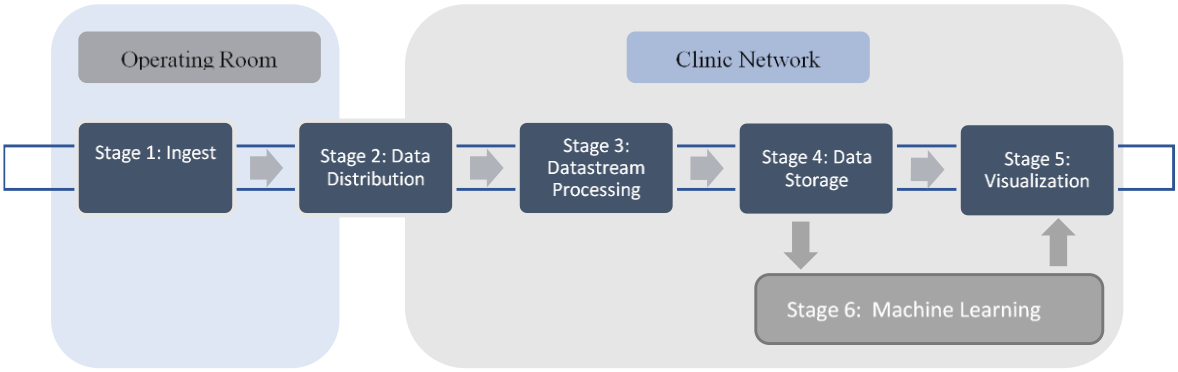


Figure 1. PART pipeline architecture with its 6 stages

III. PART PIPELINE ARCHITECTURE

As mentioned in Section II, there are several hurdles to overcome implementing a monitoring system in a clinical environment. Hence, the focus was to create a generic architecture which can be adapted and extended for future needs. This includes ingesting data from arbitrary sources, scalability, high availability, and failure safety.
In the following sections the PART pipeline architecture with its individual stages is described in detail, see Figure 1.

A. Stage 1: Ingest

Stage 1 described here is responsible for ingesting medical device data into our pipeline architecture. As a first step we connected one operating room and collected data from medical devices of our project partner Karl Storz GmbH & Co KG [11]. All devices are related to laparoscopic surgery like insufflators, endoscopic cameras, and light sources. At Heidelberg University Hospital those devices are located on a mobile cart which makes it possible to move them between surgery rooms. When the cart is moved into a surgery theater and plugged in, all devices start up and start sending records of data in 2 second intervals out of the surgery room to a specific Karl Storz machine, called Interface Control, over the Storz Communication Bus (SCB).

B. Stage 2: Data Distribution

For broadcasting the device data within the PART pipeline architecture, the distributed streaming platform Apache Kafka [12] is used, see Figure 1. It can store huge amounts of records in a fault-tolerant durable way and processes streams as they occur. Apache Kafka has three main components which are producers, brokers, and consumers. It is comparable to a message queue or an enterprise messaging system.
The interface control, see Section A, sends the device messages from the operating room via the serial interface RS232 to a machine in a technical room, located close to the surgery theater. On this machine the records of machine data are transformed into Fast Healthcare Interoperability Resources (FHIR) [13] formatted JSON [14] objects.
FHIR is a standard describing data formats and elements and an API for exchanging electronic health records created by HL7. One of its goals is to ease the interoperation between health care systems. It provides automatic and detailed electronic data capture of operational device data and offers data formats such as JSON, XML, and RDF.

Those objects are published as streams of records by the Kafka producer via TCP and Kafka protocol for subscription by consumers in the clinic network, see Figure 1. The streams of records are grouped in categories called topics. Each record consists of a key, a value, and a timestamp.
As shown in Table 1 currently 12 topics exist from four different devices including endoscopy camera (endocam), light source, operation room light and insufflator.

Table 1. Device topics consumed and processed in the PART pipeline architecture

Topic	***Unit***
INSUFFLATOR ACTUAL FLOW	liter per minute (l/min)
INSUFFLATOR ACTUAL PRESSURE	millimeter Mercury column (mm [Hg])
INSUFFLATOR INSUFFLATION ON	0x00 off 0x01 on
INSUFFLATOR GAS VOLUME	milliliter (ml)

Topic	*Unit*
INSUFFLATOR TARGET FLOW	liter per minute (l/min)
LIGHT SOURCE INTENSITY	percentage (%) [0-100]
LIGHT SOURCE STANDBY	0x00 off 0x01 on
ENDOCAM BRIGHTNESS	Low, Medium, High, Peak, Small Scope A, Small Scope B
ENDOCAM ENHANCEMENT	Off, Low, High, Fiberscope Filter A, Fiberscope Filter B
ENDOCAM SHUTTER	Auto, 1/50, 1/85, 1/125, 1/175,1/250, 1/350, 1/500, 1/700, 1/1000, 1/1500, 1/2100, 1/2800, 1/4000, 1/6000, 1/8500, 1/12000,1/17000
ORLIGHT1 INTENSITY	percentage (%) [0-100]
ORLIGHT2 INTENSITY	percentage (%) [0-100]

C. Stage 3: Data Stream Processing

As operator, the Heidelberg University Hospital wants to know when a device is used, how often it is used, and what is the ratio between hours of operation and device utilization. Those numbers help the management planning device replacements and optimizing the procurement process of the medical equipment.

As stated in Section III.B Apache Kafka is used to handle the vast amount of streams of records coming from the medical devices out of the surgery theater. As the next stage in the PART pipeline architecture Apache Flink [15] then consumes and analyses those streams of records in real-time.

The data stream processing stage, described here, focuses on specific continuous streams of records of the Karl Storz devices, which indicate whether the device is online or in standby, see Section III.B. Each record of those streams contains either a value equal to zero (standby) or equal to one (online). Hence, a device is switched on when the values of the records turn from zero to one and is offline if no data is sent from the device at all.

For all medical devices currently connected to the system, an Apache Flink program, called job, is implemented analyzing the incoming data streams in real-time.

All jobs are organized and managed as a standalone cluster on a machine in the clinic network and working as consumers subscribing to the specific topics sent by the Kafka producer to the Kafka broker. Each job can be separated into several steps, described subsequently.

For each stream:

- Transform
  Each incoming JSON structured record is disassembled and mapped onto a Flink container structure, which keeps all relevant parameters for the analytics.
- Split
  After the mapping, the new stream is split into a main stream (depicted in Figure 2 as the blue graph), containing all records of data and a sub stream (shown as the red curve in Figure 2), containing records of values equal to one. Note that the split of the sub stream is only done when the device is switched on.
- Process
  To work on infinite streams, Flink offers the concept of windows [16]. Session windows split the stream into chunks of data with finite size by sessions of activity. This is the case every time when a device is plugged in, and data is ingested into the pipeline architecture. In Figure 2 there are two session windows for the main stream, colored in blue, and three session windows for

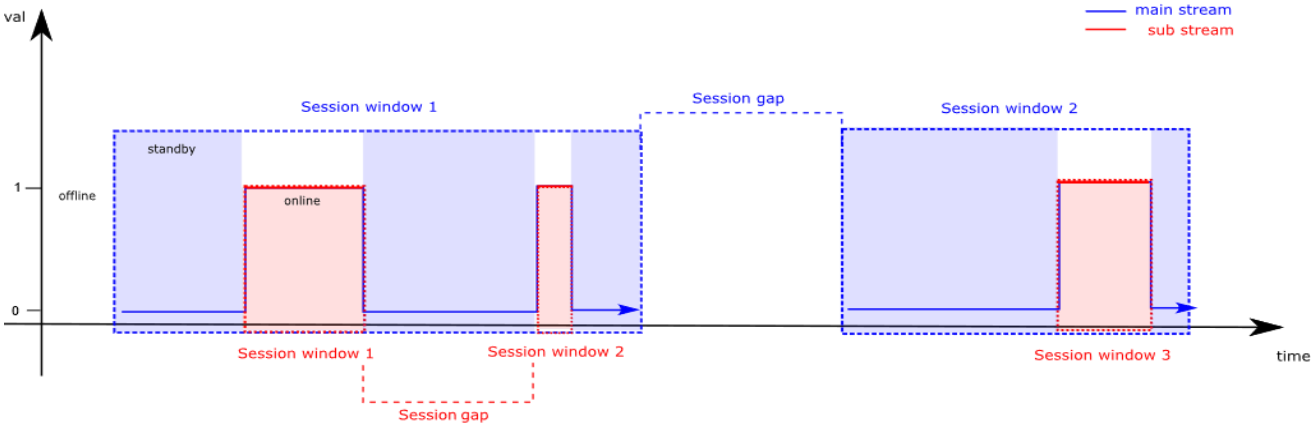


Figure 2. Data stream processing. The original stream is split into two streams. The main data stream colored in blue contains all incoming records. The sub stream colored in red contains records with values equal 1. Each partitioned into session windows separated by a predefined session gap. The processing is done on each individual window.

Figure 3. Insufflator dashboard. Key performance indicators of the observed device within a two-month interval

the sub stream, colored in red. Those windows do not have a fixed start or end. Flink closes those windows after a period of inactivity and assigns subsequently coming events to the next window. This period can be a predefined fixed interval or dynamically extracted and is called session gap.

To implement session windows, it is required to conduct the concept of event time [17] in the Flink application. This guarantees that all incoming events are ordered time wise and are processed when they happen. Even with delays, e.g., network traffic or how fast the stream is processed by the application, it yields to deterministic results. Contrary to this, the concept of Flink processing time would mean processing the stream of records when they receive the application.

Once Flink closes a session window, a process function is applied on that finite stream of data. It computes the start, end, and the duration time for the sub stream, indicating the time the device is online (depicted in Figure 2 red graph) and the values start, end, and operation time of the main stream (depicted in Figure 2 blue graph), indicating the whole period of time the device is plugged in and sending data. Note that the duration time of a device is always equal or less then the operation time.

Those results are then issued as new events separately, narrowing down an infinite stream of data, on the relevant information we were looking for.

- Store

  The results then are mapped back from the Apache Flink data structure into a valid JSON object, required by Elasticsearch [18] for storage in the PART pipeline architecture, see Figure 1. A Flink connector is used to create an Elasticsearch sink, writing the results via REST to the PART data store.

D. Stage 4: Data Storage

As stated in Section I, the monitoring system for networked medical devices at Heidelberg University has several requirements according to a data store. In the presented architecture medical equipment of Karl Storz is sending JSON formatted data via Kafka protocol to the database. However, to handle other types of data from different manufacturers, e.g., not related to surgery theaters, there is a need for a flexible system, which offers tools for ingestion of arbitrary sources of data. Connecting more of those devices gradually will yield to surging amounts of data. The data store must have the ability to adapt to changing conditions as needed overcoming problems of access times and failure safety.

The Elastic Stack offers open-source solutions for those mentioned requirements. For ingesting data there are lightweight data shippers called Beats [19] and the data collection engine Logstash[20]. Storage, search, and simple analytics is done by Elasticsearch, which runs as a cluster and scales horizontally. It is capable storing complex data structures represented as JSON objects by using RESTful APIs [21].

In the PART pipeline architecture, all observed data streams and the results of the real-time analytics are stored in Elasticsearch into an index. An index is comparable to a data table in the concept of relational database systems (RDBMS). For the index, an underlaying schema handles the mapping between JSON data fields and Elasticsearch

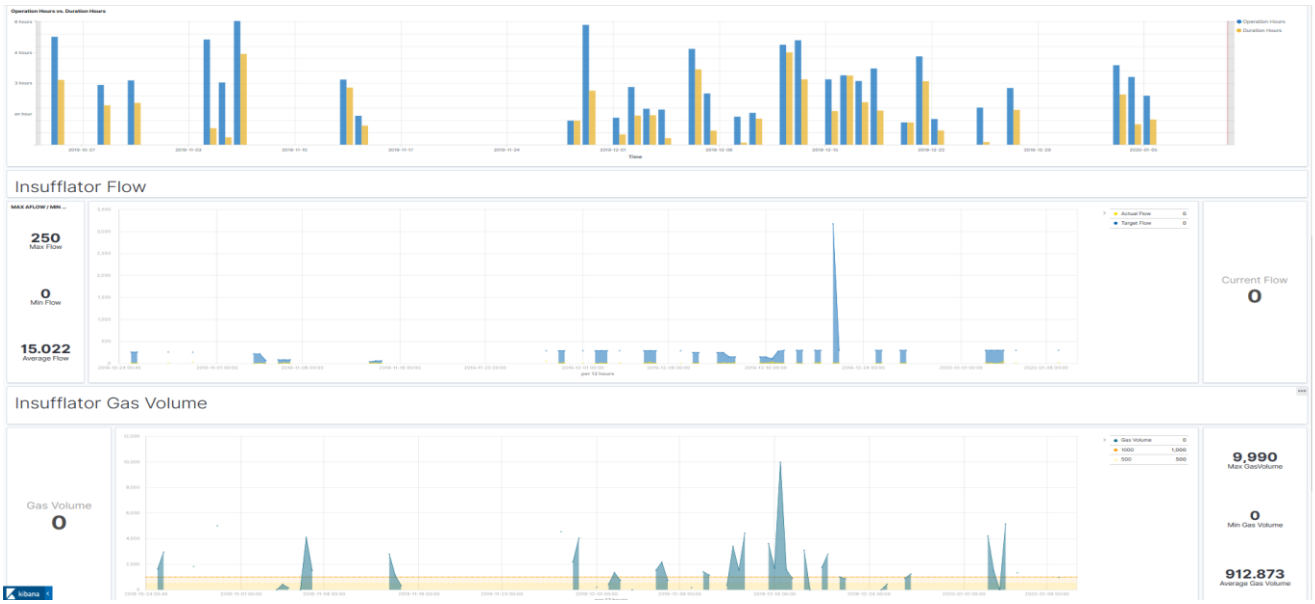


Figure 4. Insufflator dashboard. Recorded process values of the insufflator for a period of two month

data types during the ingestion. The data is then used for further analysis and visualization in stage 5.

E. Stage 5: Visualization

Kibana [22] is a data exploring and visualization tool and runs on top of Elasticsearch. It is part of the Elastic Stack and helps interacting with huge amounts of stored sensor data. It provides tools for searching, analyzing, and presentation and is used in a wide field of applications, e.g., Industrial Internet of Things (IIoT) [23][24].

Easy understandable, clearly structured visualizations embedded into good organized dashboards help conceiving the data more quickly and give first insights and hints about potential problems, e.g., device failures, even without applying sophisticated AI algorithms. Further, such dashboards help keeping track of a growing number of networked medical devices and increase the visibility and status of each individual one. Hence, Kibana is used in our PART pipeline architecture for descriptive analytics by creating dashboards combining simple histograms, line graphs, and more advanced time series aggregations for the monitored medical equipment.

F. Stage 6: Machine Learning

The problems with technical equipment during laparoscopic surgeries have been analyzed in early years which show us that there were different issues accordingly [25]. We aim to carry the current situation to a next step which would be beneficial to implement machine learning model on device data to be able to decrease the technical problems during the surgeries.

In machine learning, understanding the data is a key. Therefore, Jupyter Notebook [26] is used to analyze the data for machine learning implementation. In this phase of the architecture, the data from Elasticsearch is taken to Jupyter Notebook by using Pandasticsearch library [27]. In this way, we can create Pandas data frame for data analysis as well as implement a machine learning algorithm.

To be able to train machine learning algorithms, the data should be separated into a clean, annotated, well-structured dataset. These datasets then will be stored in MongoDB [28] as train, test, and validation datasets separately. With that, we can train our machine learning algorithm on training datasets and test the accuracy on test datasets. The aim of implementing machine learning algorithms will be used for the use case 2 “online inspection”, to find the anomalies of the devices through their data. Therefore, we assume that unsupervised machine learning techniques will fit the most to find anomalies and outliers.

IV. DASHBOARD

Our goal was to create an intuitive dashboard for the monitoring tool, including all medical devices in the surgery theater. As stated in Section III.E we use Kibana for

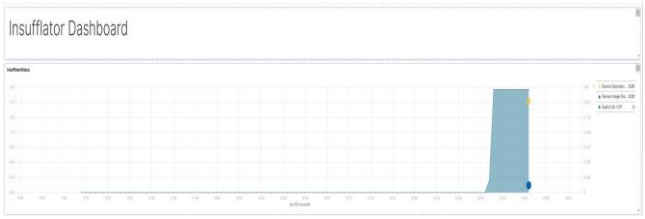


Figure 5. Insufflator Status. Device is offline, in standby or online

visualization of laparoscopy device data. For each device type, e.g., insufflators, one monitoring dashboard was created by assembling different kinds of visualizations.
Subsequently the PART dashboard (see Figure 3, Figure 4) with focus of profitability analysis and system monitoring for the insufflator is shown.
The dashboard has a date range field which enables looking at specific time ranges getting different insights of the key performance indicators (KPI). This makes it possible to examine historical periods of time, e.g., the last two years, shorter periods, e.g., minutes or even real-time data.

The presented dashboard is divided into two sections. Figure 3 shows the top section, including visualizations and numbers according to the profitability analysis, and the system monitoring for a time period of two month exemplarily. Followed by a section showing medical device data related to the process values, including the insufflator pressure, flow and the gas volume for the same time range, depicted in Figure 4. Subsequently, some of the visualizations and graphs of the dashboard are described in more detail. Figure 5 shows the top section of the dashboard which is the insufflator status. It plots the stream of records grouped by the topic INSUFFLATION ON, see Table 1. Note that in Figure 3 and Figure 4 a much wider range of time is selected, hence more data is visualized. In Figure 5 we basically zoomed into the data timewise, showing a two-hour period of recorded insufflator data for better understanding. The graph shows the usage of the insufflator during that time period. Three scenarios are possible:

- offline - no data is available
- standby - the device is sending records of zero data
- online - the device sending records with values of one

The x-axis shows the time, the y-axis has two scales. On the left-hand side it shows the status, zero for standby, one for online. On the right it shows the time in hours according to the device usage and operation time.
If there is no data shown on the timeline, which is the case at the very beginning and the end of the graph, the device is switched off. In the time period where the graph shows values of zero, the device is in standby. This is the situation usually right before a surgery, where the cart is already moved into the surgery room and plugged in, see Section III.A. When the graph values turn to one, it indicates that the device is now being used. After a short period of time it was switched off again. In this visualization a laparoscopic intervention is shown.
When switched off, Apache Flink, see Section III.C, calculates the operation and the usage time indicated by the yellow and the blue dot (see Figure 5). Operation time

Figure 6. Insufflator KPIs – duration time, operation time and usage counter

implies the whole time the device is powered on. Usage time represents the time the device is actually used. The difference between those measures is encoded in their height. In our example there is a long time where the device is in operation but only a small period of time where it is actually used. From an economic perspective and as the operator of thousands of devices at Heidelberg University Hospital those numbers are analyzed according to potential optimization.
In the section below, the status graph (see Figure 6) of the PART dashboard some KPIs, e.g., the device usage- and operation time of the insufflator, are shown. As depicted in Figure 6 the very left panel shows the usage of the device, in our case eight minutes, and the total operation time on the right, which is about two hours. In the middle of this section a panel shows the time when the device was online the last time, named as last seen, and a usage counter for the specific time frame. Note that this information can be exported as a report.
To give the user an overview about those numbers on a daily basis, a calendar was created for usage and operation time, see Figure 7. The heat map visualization of Kibana does a color coding from white to blue according to the duration, where blue indicates a longer usage duration. This utilization distribution immediately shows peaks and helps finding patterns in device and operating room usage, which can help optimizing utilization of medical equipment by balancing the product use and hence could yield to less inventory.

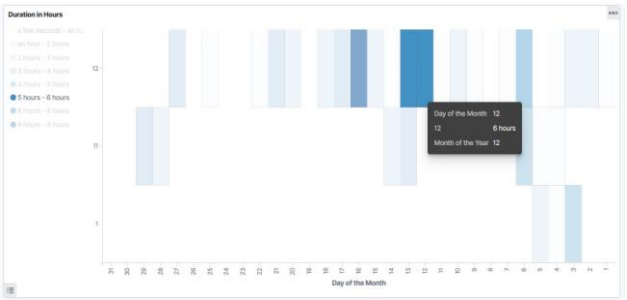


Figure 7. Calendar Visualization: Heatmap of the device usage per day. Dark blue indicates longer usage times

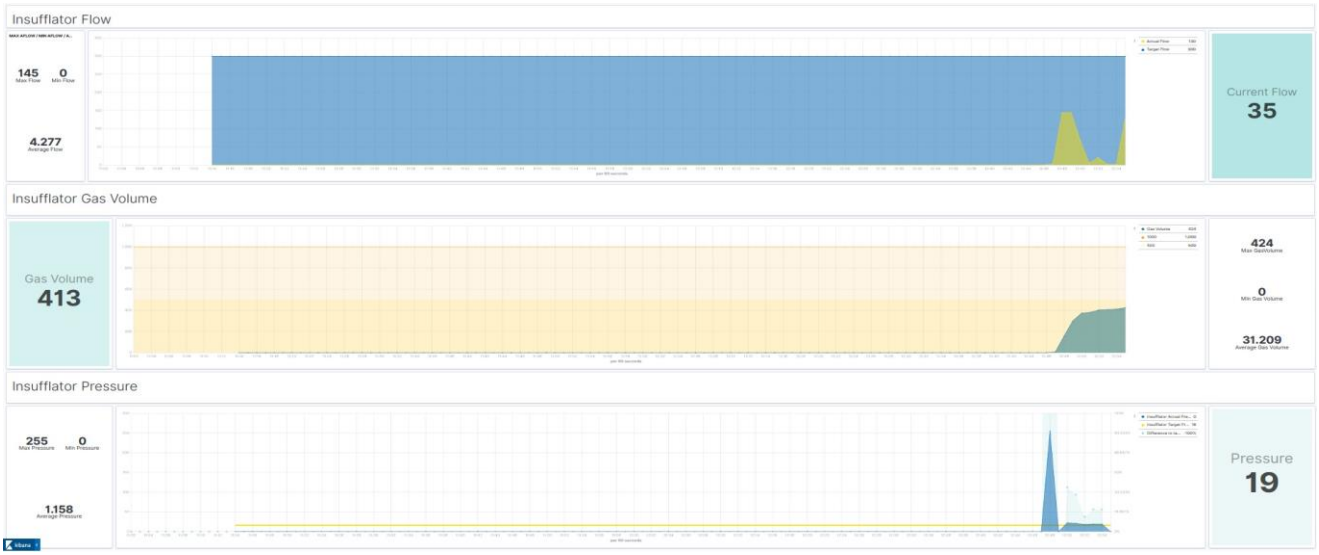


Figure 8. Detailed insufflator process values during a surgery. Flow, gas volume and pressure visualizations in the PART dashboard

The lower part of the insufflator dashboard has three panels each divided into three sections, see Figure 8. The middle section of each panel contains a graph drawing the process values. On the x-axis of this graph the time is indicated. On the far right, the graph shows the latest record of the data stream received by the monitoring tool. There is one section depicting minimum, maximum, and the average of the process values. One field shows the latest value recorded in Elasticsearch as a number, which is identical to the far-right value of the graph.

At the top of Figure 8 the insufflator flow is shown. The blue area depicts the target, the yellow graph the actual flow. Currently, the latter has the value 35. The progress over time of the insufflator gas volume is depicted in the middle. The accumulated value is 413. At the bottom, the user can monitor the pressure of the insufflator. The yellow line indicates the target pressure while the blue curve represents the actual pressure, currently set to 19, during a surgery. Those visualizations enable the user to monitor the devices and the whole surgery theater as being used in real-time detecting potential anomalies in the data only by descriptive analytics.

V. Discussion and Outlook

The goal of PART is to build an AI driven monitoring system for networked medical devices. This system should be vendor independent. As shown in Section II, there are several hurdles to overcome reaching this goal. Those circumstances made us focus building up a generic architecture which is a flexible, easy to extend infrastructure for ingesting, analyzing, storing, and visualizing medical device data. We started ingesting data out of one surgery room from laparoscopy related devices of one device manufacturer. We are analyzing those streams of records in real-time according to the use case profitability analysis and built a dashboard showing those results for the monitored medical devices.

Still, at Heidelberg University Hospital, our strategy is to add more devices of different manufacturers gradually. Hence, we work closely together with other device manufacturers, not only related to operating rooms, e.g., patient monitoring.

Further, we continue to work on the architecture to be prepared for growing amounts of medical device data in the future. This is done by moving the PART pipeline architecture into the Medical Data Integration Center (MeDIC) of the Heidelberg University Hospital. Including the transformation of each particular stage of the PART pipeline architecture from standalone components to orchestrated clusters. Making the monitoring system more robust against performance issues and failure safety. Additionally, a monitoring system for the PART pipeline architecture with its components should be implemented to handle the complexity of the monitoring system itself.

Another upcoming task is to connect more integrated operating theaters [15] with its laparoscopic devices from Karl Storz including collection, analyzing, and visualizing the retrieved data as presented here. Those extensions and the goal to support a multitude of devices from different vendors, requires a complete redesign of the PART dashboard in the future.

There is still a lot of information in the data which is not appropriately extracted and visualized. Hence, extending the dashboards by additional visualizations is planned. For example, including distribution numbers when specific devices are used during a day, month or year or adding a counter which shows since when a device is online.

With the experience made, we have lowered our aspiration towards an AI driven system for predictive maintenance and hence focused on a combination of descriptive analytics and

real-time stream processing as a first approach. However, we still have the ambition to push this topic further by extending the PART pipeline architecture with machine learning tools and deepen the research on the other use cases. Therefore, we look at the data comprehensively to be able to create a well-structured dataset that can be used straightforwardly for further implementations. To be able to succeed this, we work with surgeons closely to annotate data.

Establishing such a generic monitoring system at Heidelberg University Hospital would have several benefits for the clinic. It should make the complete IT infrastructure more robust and stable. Detecting problems of networked devices in early stages, by real-time alerts or in a best-case scenario before they are going to happen, saves maintenance time and costs. Further, with workload statistics of the devices one gets a good tool for tracking the usage and can adapt the inventory accordingly.

VI. Acknowledgment

This work is funded by the German Federal Ministry of Education and Research under reference number 16KIS0785 as part of the PART project. We thank all partners within the PART project specifically Karl Storz GmbH & Co. KG supporting us with their knowledge and medical equipment.

VII. References

[1] O. Klar, R. Dees, G. Schneider, A. R. Poyraz and O. Heinze, “First Experiences Implementing Predictive Analytics Tools in a Clinical Routine Setting” in Proc. eTELEMED 2019, The Eleventh International Conference on eHealth, Telemedicine, and Social Medicine, pp. 114-115.

[2] T. Carvahlo, F. Soares, R. Vita, R. d. P. Francisco, J. P. Basto and S. G. S. Alcala, "A systematic literature review of machine learning methods applied to predictive maintenance," Computers & Industrial Engineering, 2019.

[3] R. H. Choplin, J. M. Boehme and C. D. Maynard, "Picture archiving and communication systems: an overview." RadioGraphics 12(1): pp. 127-129, 1992.

[4] Draeger, Medical device connectivity based on IEEE 11073 SDC. [Online]. Available from: https://www.draeger.com/en_me/Hospital/Acute-Care-Insights/Service-Oriented-Device-Connectivity 2020.01.28

[5] Philips e-Alert. [Online]. Available from: https://www.philips.co.uk/healthcare/product/HC895000/philips-ealert-alerting-solution-for-mri-systems 2020.02.12

[6] R. B. Patil, M. A. Patil, V. Ravi and S. Naik, "Predictive modeling for corrective maintenance of imaging devices from machine logs," 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Seogwipo, 2017, pp. 1676-1679.

[7] Siemens Teamplay. [Online]. Available from: https://www.siemens-healthineers.com/de/digital-health-solutions/digital-solutions-overview/service-line-managment-solutions/teamplay 2020.01.27

[8] Philips PerformanceBridge. [Online]. Available from: https://www.philips.de/healthcare/services/performance-services/performancebridge 2020.01.27

[9] Stryker Smart Equipment Management. [Online]. Available from: https://www.stryker.com/us/en/surgical/services/smart-equipment-management.html 2020.01.27

[10] V. Gudivada, A. Apon, and J. Ding. ”Data Quality Considerations for Big Data and Machine Learning: Going Beyond Data Cleaning and Transformations". In: International Journal on Advances in Software 10.1 (2017), pp. 1 – 20

[11] Karl Storz OR1. [Online]. Available from: https://www.karlstorz.com/de/de/karl-storz-or1.htm 2020.01.27

[12] Apache Kafka. [Online]. Available from: https://kafka.apache.org/ 2020.01.27

[13] B. Franz, A. Schuler and O. Krauss. "Applying FHIR in an integrated health monitoring system." EJBI 11.2, 2015, pp. 51-56.

[14] JSON. (2020). [Online]. Available from: https://www.json.org/json-en.html 2020.01.27

[15] Apache Flink. [Online]. Available from: https://flink.apache.org/ 2020.01.27

[16] Apache Flink Windows. [Online]. Available from: https://ci.apache.org/projects/flink/flink-docs-stable/dev/stream/operators/windows.html 2020.02.04

[17] Apache Flink Event Time. [Online]. Available from: https://ci.apache.org/projects/flink/flink-docs-stable/dev/event_time.html 2020.02.04

[18] Elastic. [Online]. Available from: https://www.elastic.co/de/ 2020.01.27

[19] Elasticsearch Beats. [Online]. Available from: https://www.elastic.co/de/beats 2020.02.06

[20] Elasticsearch Logstash. [Online]. Available from: https://www.elastic.co/de/logstash 2020.02.06

[21] R. T. Fielding, "Architectural Styles and the Design of Network-based Software Architectures," [Online]. Available: https://www.ics.uci.edu/~fielding/pubs/dissertation/fielding_dissertation_2up.pdf. [Accessed 29 1 2020].

[22] Kibana, (2020). [Online]. Available from: https://www.elastic.co/de/kibana 2020.01.27

[23] IIoT with the Elastic Stack. [Online]. Available from: https://www.elastic.co/de/blog/industrial-internet-of-things-iiot-with-the-elastic-stack 2020.01.27

[24] M. Bajer, ”Building an IoT Data Hub with Elasticsearch, Logstash and Kibana. *Building an IoT Data Hub with Elasticsearch, Logstash and Kibana”,* 2017.

[25] E.G.G. Verdaasdonk, L.P.S. Stassen, M. van der Elst, T. M. Karsten and J. Dankelman, “Problems with technical equipment during laparoscopic surgery. An observational study“ , Surg Endosc (2007), pp. 275-279.

[26] Jupyter Notebook. [Online]. Available from: https://jupyter.org/ 2020.01.27

[27] Pandasticsearch. [Online]. Available from: https://pandasticsearch.readthedocs.io/en/latest/index.html# 2020.01.27

[28] MongoDB. [Online]. Available from: https://www.mongodb.com/ 2020.01.27

[29] M. Kasparick et al. “New IEEE 11073 standards for interoperable, networked point-of-care Medical Devices,” in Proc. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE, 2015, pp. 1721-1724

An Information-Theoretic Model for Managing Predator–Prey Ecosystems Facing Climate Changes

Katsumi Sakata
Department of Life Science and Informatics
Maebashi Institute of Technology
Maebashi, Japan
e-mail: ksakata@maebashi-it.ac.jp

Ramesh Katam
Department of Biological Sciences
Florida A&M University
Tallahassee, USA
e-mail: ramesh.katam@famu.edu

Toshiyuki Saito
Department of Radiation Effects Research
National Institute of Radiological Sciences
Chiba, Japan
e-mail: saito.toshiyuki@qst.go.jp

Hajime Ohyanagi
Computational Bioscience Research Center
King Abdullah University of Science and Technology
Thuwal, Kingdom of Saudi Arabia
e-mail: hajime.ohyanagi@kaust.edu.sa

Setsuko Komatsu
Department of Environmental and Food Sciences
Fukui University of Technology
Fukui, Japan
e-mail: skomatsu@fukui-ut.ac.jp

***Abstract*—We have developed a theoretical model for managing predator–prey ecosystems undergoing climate-related changes. This model estimates the amount of information transferred between the number of predator and prey categories, and the uncertainty of the number of the prey categories in a predator–prey ecosystem as measured by Shannon entropy, which is achieved by predation events and decay in the ecosystem. We examined the model with a numerical experiment using a well-studied bass–crayfish predator–prey ecosystem in a closed lake. Furthermore, we have evaluated the model comparison using Lotka-Volterra equation, which is a conventional predator–prey ecosystem model.**

Keywords-Mathematical model; Information theory; Predator–prey ecosystem; Climate change.

I. INTRODUCTION

This article is a revised version of conference paper presented at the Eleventh International Conference on Bioinformatics, Biocomputational Systems and Biotechnologies [1]. This extended version of the original paper offers results from our preliminary studies on a mathematical model for predator–prey ecosystems facing climate changes.

Over the past three decades, environmental changes, such as global warming, desertification, and air pollution have become much worse, raising the concerns of their effects on life systems [2]. Previous studies on the environmental responses of life systems focused on specific networks such as the genetic and ecological networks [3,4].

Life system takes orderliness from its environment and sustains itself at a fairly high level of orderliness, or at a fairly low level of thermodynamic entropy [5]. Kauffman investigated on how the dynamic behavior of a Boolean network suddenly becomes orderly. He made the analogy that the behavior approximates cell fate which is characterized by expression patterns of multiple genes in an organism [6,7]. Barabási and Albert found that generic mechanisms form an ordered network structure with a scale-free property [8]. However, the previous observations were not validated by a mathematical model that clarifies the varying orderliness of biological systems undergoing environmental changes.

Recent studies for mathematical model of predator-prey ecosystems, examined the model with double free boundaries [9], developed an agent-based model of an ecosystem to predict interactions of competition and predation [10]. Further studies showed a simple two-species predator-prey ecosystem that can display rich dynamical complexity when the prey evolves in response to predation, based on coupled differential equations [11] .

Environmental factors including air pollution could influence the prey-predators relation [12,13]. Atmospheric change may influence predator-prey interactions by altering prey quality, defensive behavior of prey, predator location, prey community structure and/or predator competition [12]. Fish behavior can be altered by contaminants as a means that the mummichog from more contaminated areas are poor predators and slower to capture active prey, the grass shrimp [13]. Thus, to study the simultaneous effect of pollution stress and the effect of infection in an interacting species is

important for deriving the feasible situations of an ecosystem [14].

In this study, we quantify the environmental stimuli and orderliness achieved in state variables in life systems with Shannon entropy based on their probability distributions. The state variables represent the state of the system, such as expression levels in a genetic network. We then hypothesize a relationship between environmental changes and orderliness in the life systems. We validate the hypothetical relationship using numerical experiments based on a computational model of differential equations for the ecosystem with the climate-shift model [15]. In the model, a climate-attribute change is modeled as a shift in the probability distribution of the climate attribute. We evaluate control performance by a difference of Shannon entropy as $\Delta H \equiv H(X) - H(X')$, where X and X' represent the state variable X at t_0 and at t_1 (unit time after t_0), respectively [16]. The Shannon entropy $H(X)$ indicates the uncertainty of X [17]. We elucidate predator–prey ecosystem degradation, which is unavoidable, in the current progression of climatic changes and develop an information-theoretic framework for performing systematic countermeasures against climate change. The details are given below under Section II: formulation, Section III: numerical experiments, and Section IV: comparison with Lotka-Volterra equations. Finally, in Section V, we conclude with final remarks and discuss future works.

II. FORMULATION

A closed lake is considered in our study as a predator–prey ecosystem (Figure 1a). The probability distribution of the number of viable predators, known as "capacity", varies according to the climate shift of a climate attribute against a range of climate attributes (survival region) in which the predator is viable. The predator capacity decreases with an increase in climate shift (Figure 1b). We derived Equation (1), which shows that the Shannon entropy ($H(Y)$) of the number of predators decreases with an increase in the climate shift (Figure 1c) based on a numerical analysis (see "Numerical Analysis for Equation (1)" at the end of this section):

$$H(Y)_{e+\delta e} \le H(Y)_e , \tag{1}$$

where e and δe indicate the level of the climate attribute and its increment. Generally, $I(X;Y) \le \min\{H(X), H(Y)\}$, thus

$$I(X;Y)^U_{e+\delta e} \le I(X;Y)^U_e , \tag{2}$$

where $I(X;Y)^U (\equiv H(Y))$ denotes an upper bound of the mutual information between X and Y. The causal relationship in the ecosystem model shown in Figure 1a is derived and shown as in Figure 2. Predators limit the growth of prey by consuming the prey [18]. Increases or decreases in the prey population lead to corresponding increase or decrease in the number of predators. This scenario configures a closed-loop circuit into which an environmental stimulus can be introduced. An information–theoretic limit (3) for closed-loop control systems [16]:

$$\Delta H_{\text{closed}} \le \Delta H^{\max}_{\text{open}} + I(X;Y), \tag{3}$$

where ΔH_{closed} and $\Delta H^{\max}_{\text{open}}$ are the Shannon entropy reduction of the state variable X in a closed-loop control system and the maximum Shannon entropy reduction of the state variable X in a general open-loop control system, respectively, over the transition $X \to X'$ between t_0 and t_1 (unit time after t_0). We merged Equation (2) with the

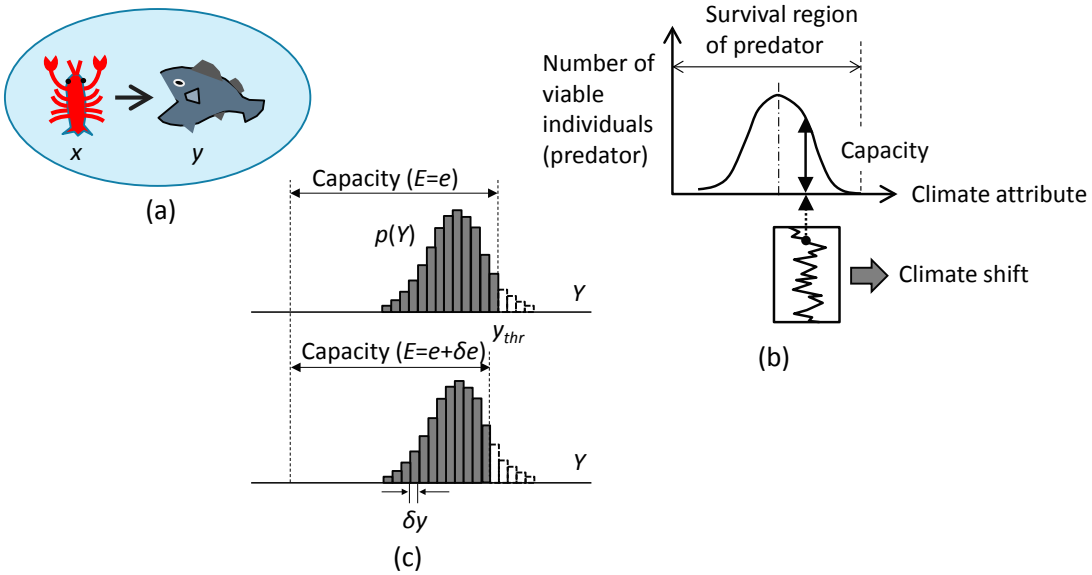


Figure 1. An ecosystem model: (a) Predator–prey ecosystem in a closed lake. The arrow denotes feeding relationship. (b) Number of viable predators and climate shift. (c) Probability distribution of the number of predators before (upper panel) and after (lower panel) an increase in climate shift by δe.

information–theoretic limit for closed-loop control systems (Equation (3)), to derive Equation (4):

$$\Delta H^{U}_{e+\delta e} \leq \Delta H^{U}_{e}, \quad (4)$$

where $\Delta H^{U}(\equiv \Delta H^{\max}_{\text{open}} + I(X;Y)^{U})$ denotes an upper bound of the Shannon entropy reduction of the state variable X (the number of prey individuals) over the transition $X \rightarrow X'$ between t_0 and t_1. It represents the control performance of the predator–prey ecosystem. Equations (2) and (4) suggest that the mutual information between the number of prey individuals (X) and predators (Y), as well as the control performance of the predator–prey ecosystem, decreases with an increase in climate shift. Furthermore, the control performance of the predator–prey ecosystem appears to degrade from the level of a closed-loop control system to an open-loop control system, based on the information–theoretic limits of control (Equation (3)).

The derived inequalities of Equations (1), (2) and (4) are independent of the dynamics of the target ecosystem. Thus, our model can be applied to analyses of unknown dynamics in ecosystems.

It should be noted that the decrease in the reduction of uncertainty in the state variable (represented by ΔH^{U} in Equation (4)) is unavoidable under the progressive environmental change ($\delta e \geq 0$).

Numerical Analysis for Equation (1)

We calculated the Shannon entropy of the number of predators $H(Y)$ against the climate shift (Figure 3). We assumed that Y (number of predators) follows a normal distribution, and calculated a normalized $H(Y)$. The normalized $H(Y)$ =1 and 0 indicate that the Shannon entropy $H(Y)$ is equivalent to the Shannon entropy at E=0 (the region between $\mu-3\sigma$ and $\mu+3\sigma$ of the normal distribution lies inside of the capacity of the predator) and 1 (the region between $\mu-3\sigma$ and $\mu+3\sigma$ of the normal distribution gets out of the capacity of the predator), respectively. Where, μ and σ are the mean and the standard deviation of the normal distribution of the number of predators, respectively. As the result the normalized $H(Y)$ continuously decreases (Figure 3).

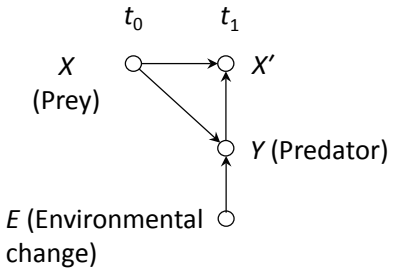


Figure 2. Causal relationship in the ecosystem model represented in a directed acyclic graph [19].

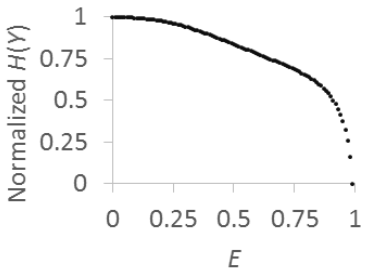


Figure 3. Normalized $H(Y)$ against the climate shift.

III. Numerical Experiments

Equations (2) and (4) suggest the two relationships: mutual information between the number of prey and predator individuals decreases with an increased climate shift (relationship-1). Control of the number of prey individuals becomes worse in a predator–prey ecosystem with an increased climate shift (relationship-2).

A. Numerical Experiment Methods

To validate these two relationships and evaluate management actions for protecting predator–prey ecosystems against climate-related changes, we conducted numerical experiments in which we used a well-studied nondimensionalized bass–crayfish predator–prey ecosystem in a closed lake [20]:

$$\frac{dx}{dt} = x(1-x-\alpha y) - \frac{\delta y x^2}{\kappa^2 + x^2}$$
$$\frac{dy}{dt} = \gamma y(1-\beta x - y) - \varepsilon \frac{\delta y x^2}{\kappa^2 + x^2} \quad (5)$$

where $x = X/K_x$, $y = Y/K_y$ (X and Y: crayfish and bass biomass, K_x and K_y: crayfish and bass carrying capacity), and α, β, δ, γ, ε, and κ are model parameters (positive real constants), and were set to 0.7, 0.9, 0.075, 1.5, 0.01, and 0.1, respectively. Random noise with normal distribution with mean=0, and variation=0.02^2 was added to x and y. The right side of Equation (5) represents Lotka–Volterra-style intra- and interspecific competition and predation [4].

B. Validation of the Degradation of the Mutual Information and the Control Performance Suggested by the Model

The nondimensionalized biomass of bass y was limited to mimic the decreased capacity by the increasing climate shift. The decrease of the mutual information between the nondimensionalized biomass of crayfish x and bass y by the decreasing capacity for the nondimensionalized biomass of bass was observed (Figure 4). The result suggests the relationship-1 is valid. The increase of the control bias by the decreasing capacity for the nondimensionalized biomass of bass was observed (Figure 5). This result suggests the relationship-2 is valid.

C. Evaluation of the Management Actions for Protecting Predator–Prey Ecosystems Against Climate-Related Changes

From the definitions $I(X;Y)^U \equiv H(Y)$ and $\Delta H^U \equiv \Delta H_{\text{open}}^{\max} + I(X;Y)^U$ in Formulation, ΔH^U depends on Y but independent of X.

We mimic two types of management actions, (i) reduce the number of crayfish, and (ii) increase the number of bass, and conducted numerical experiments using Equation (5) to compare the efficiency by the two management actions (Figure 6a–6c). The deviation from the average without limitation of the biomass of bass was improved with the management action for bass (Figure 6c) better than with the management action for crayfish (Figure 6b). The proposition mentioned in the previous paragraph, which is based on our model, validates the numerical experiment results.

IV. Comparison with Lotka-Volterra Equations

The Lotka–Volterra equations are a pair of first-order nonlinear differential equations, frequently used to describe the dynamics of predator-prey ecosystems [21,22,23]. Assumptions made in the creation of the Lotka-Volterra equations include:

(i) There is no shortage of food for the prey population.
(ii) The amount of food supplied to the prey is directly related to the size of the prey population.
(iii) The rate of change of population is directly proportional to its size.
(iv) The environment is constant and genetic adaptation is inconsequential.
(v) Predators will never stop eating.

Under the assumptions, the Lotka-Volterra equations are written as:

$$\frac{dx}{dt} = \alpha x - \beta xy \qquad (6)$$
$$\frac{dy}{dt} = \delta xy - \gamma y$$

where x and y are the number of prey and predators, respectively, and α, β, δ, γ are model parameters (positive real constants) [24]. As the differential equations are used (Equation (6)), it is implied that births, deaths, and movements are continuous, and there are overlapping generation [25].

Compared to the Lotka–Volterra equations, our model (Figure 2) only assumes the number of prey and predator interact. The continuity about the births, deaths, and movements is similarly implied in our model. The important thing is that our model allows the effect of the environmental change by the capacity for the number of predator (Figure 2).

V. Conclusions and Future Works

We have developed an information–theoretic predator–prey ecosystem model that is independent of the dynamics of the ecosystem and validated the model through numerical experiments. Numerical experiment results also suggested our model is effective for evaluating management actions for predator-prey ecosystems against environmental changes, which include the uncertainties of the environmental factors such as air pollution. We compared our model with Lotka–Volterra equations and clarified our model that needs a few assumptions but, nevertheless, still adequately predicts environment-related changes in predator-prey ecosystems. The information-theoretic framework will be useful to the environmental responses of other life systems such as genetic regulatory networks. Studies on more comparisons between the environmental responses of the life systems and elucidation of universal rules over them are needed to validate the theory.

Acknowledgment

This study was supported by the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (Grant No. 16K00399). We thank Natalie Kim, PhD, from Edanz Group (www.edanzediting.com/ac) for editing a draft of this extended abstract. We thank Takahiro Morita from Maebashi Institute of Technology for supporting literature survey. We thank Virginia Gottschalk, PhD, from Florida A&M University for helpful comments on this manuscript.

References

[1] K. Sakata, T. Saito, H. Ohyanagi, and S. Komatsu, “A Mathematical Model for Predator–Prey Ecosystems Facing Climate Changes,” in Proceedings of the 11th Intl. Conf. on Bioinformatics, Biocomputational Systems and Biotechnologies (BIOTECHNO 2019) IARIA, Jun. 2019, pp. 28–29, ISBN: 978-1-61208-717-7.

[2] T. Stocker et al. “The Physical Science Basis – Summary for Policymakers,” IPCC WGI AR5 (Report), p. 4, 2013.

[3] G. Cramer, K. Urano, S. Delrot, M. Pezzotti, and K. Shinozaki, “Effects of Abiotic Stress on Plants: A Systems Biology Perspective,” BMC Plant Biol., vol. 11, p. 163, 2011, doi: 10.1186/1471-2229-11-163.

[4] R. Horan, E. Fenichel, K. Drury, and D. Lodge, “Managing Ecological Thresholds in Coupled Environmental-Human Systems,” PNAS, vol. 108, pp. 7333–7338, May 2011, doi: 10.1073/pnas.1005431108.

[5] E. Schrödinger, “What is Life?: With Mind and Matter and Autobiographical Sketches,” Cambridge University Press, 1992.

[6] S. Kauffman, “The Origins of Order: Self-Organization and Selection in Evolution,” Oxford University Press, June 1993.

[7] F. H. Westhoff, B. Yarbrough, and R. Yarbrough, “Complexity, Organization, and Stuart Kauffman’s the Origins of Order,” J. Econ. Behav. Organ., vol. 29, pp. 1–25, 1996, doi: 10.1016/0167-2681(95)00049-6.

[8] A-L. Barabási and R. Albert, “Emergence of Scaling in Random Networks,” Science, vol. 286, pp. 509–512, Oct. 1999, doi: 10.1126/science.286.5439.509.

[9] M. Wang and J. Zhao, “A Free Boundary Problem for a Predator-prey Model with Double Free Boundaries,” arXiv:1312.7751v3, May 2014.

[10] I. Karsai, E. Montano, and T. Schmickl, “Bottom-up Ecology: An Agent-based Model on the Interactions Between Competition and Predation,” Letters in Biomathematics, vol. 3, pp. 161–180, 2016, doi: 10.1080/23737867.2016.1217756.

[11] W. Gilpin and M. W. Feldman, “A Phase Transition Induces Chaos in a Predator-prey Ecosystem with a Dynamic Fitness Landscape,” PLoS Comput Biol., vol. 13, e1005644, Jul. 2017, doi: 10.1371/journal.pcbi.1005644.

[12] R. Matyssek, N. Clarke, P. Cudlin, T. N. Mikkelsen, J-P. Tuovinen, G. Wieser, and E. Paoletti, “Climate Change, Air Pollution and Global

Challenges: Understanding and Perspectives from Forest Research, Volume 13, 1st Edition," Elsevier, Nov. 2013.

[13] J. S. Weis and Allison Candelmo, "Pollutants and Fish Predator/Prey Behavior: a Review of Laboratory and Field Approaches," Current Zoology, vol. 58, pp. 9–20, 2012, doi: 10.1093/czoolo/58.1.9.

[14] S. Sinha, O. P. Misra, and J. Dhar, "Study of a Prey-Predator Dynamics Under the Simultaneous Effect of Toxicant and Disease," Journal of Nonlinear Science, vol. 1, pp. 102–117, Mar. 2008, doi: 10.22436/jnsa.001.02.06.

[15] Q. Schiermeier, "Climate and Weather: Extreme Measures," Nature, vol. 477, pp. 148–149, 2011, doi:10.1038/477148a.

[16] H. Touchette and S. Lloyd, "Information-Theoretic Limits of Control," Phys. Rev. Lett., vol. 84, pp. 1156–1159, 2000, doi: 10.1103/PhysRevLett.84.1156.

[17] D. Robinson, "Entropy and Uncertainty," Entropy, vol. 10, pp. 493–506, 2008, doi: 10.3390/e10040493.

[18] E. H. Nelson, C. E. Matthews, and J. A. Rosenheim, "Predators Reduce Prey Population Growth by Inducing Changes in Prey Behavior," Ecology, vol. 85, pp. 1853–1858, 2004, doi:10.1890/03-3109.

[19] T. D. Nielsen and F. V. Jensen, "Causal and Bayesian Networks. Bayesian Networks and Decision Graphs (Information Science and Statistics)," Springer, 2007.

[20] K. Drury and D. Lodge, "Using Mean First Passage Times to Quantify Equilibrium Resilience in Perturbed Intraguild Predation Systems," Theor. Ecol., vol. 2, pp. 41–51, 2009, doi: 10.1007/s12080-008-0027-z.

[21] A. A. Berryman, "The Origins and Evolution of Predator - Prey Theory," Ecology, vol. 73, 1992, doi: 10.2307/1940005.

[22] X. Liu and L. Chen, "Complex Dynamics of Holling Type II Lotka–Volterra Predator–Prey System with Impulsive Perturbations on the Predator," Chaos, Solitons & Fractals, vol. 16, pp. 311–320, 2003.

[23] X. P. Yan and Y. D. Chu, "Stability and Bifurcation Analysis for a Delayed Lotka–Volterra Predator–Prey System," J. Comput. Appl. Math., vol. 196, pp. 198–210, 2006.

[24] T. Blaszak and W. Hu, "Lotka-Volterra models of Predator-Prey Relationships," web.mst.edu, Retrieved 2019-09-30.

[25] W. W. Murdoch, C. J. Briggs, R. M. Nisbet, W. S. Gurney, and A. Stewart-Oaten, "Aggregation and Stability in Metapopulation Models," Am. Nat., vol. 140, pp. 41–58, 1992.,

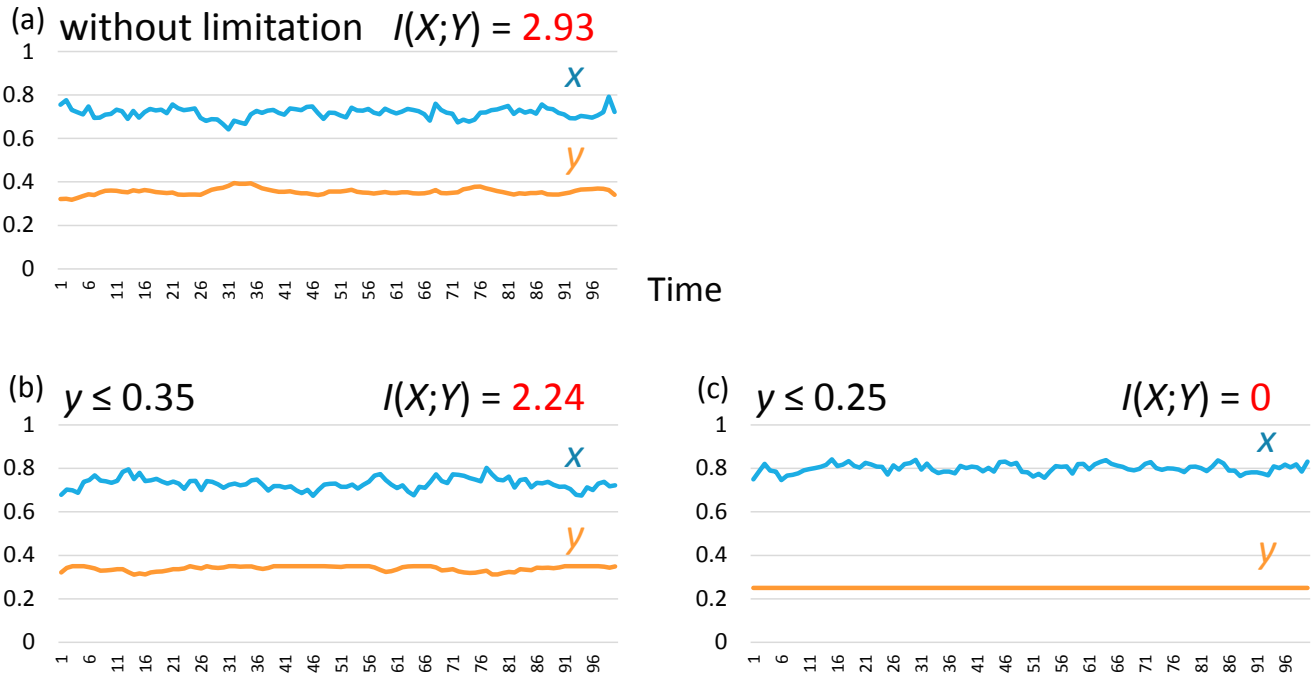


Figure 4. Numerical experiment results for the mutual information between the number of prey individuals and predators: (a) Without limitation for the nondimensionalized biomass of bass *y*. In (b) and (c), the nondimensionalized biomass of bass *y* was limited equal or lower than 0.35 and 0.25, respectively, to mimic the decreased capacity by the climate shift. Horizontal axes mean the number of iteration in the calculation for the differential equations (Equation (5)). $I(X;Y)$ means the mutual information between the nondimensionalized biomass of crayfish *x* and bass *y*.

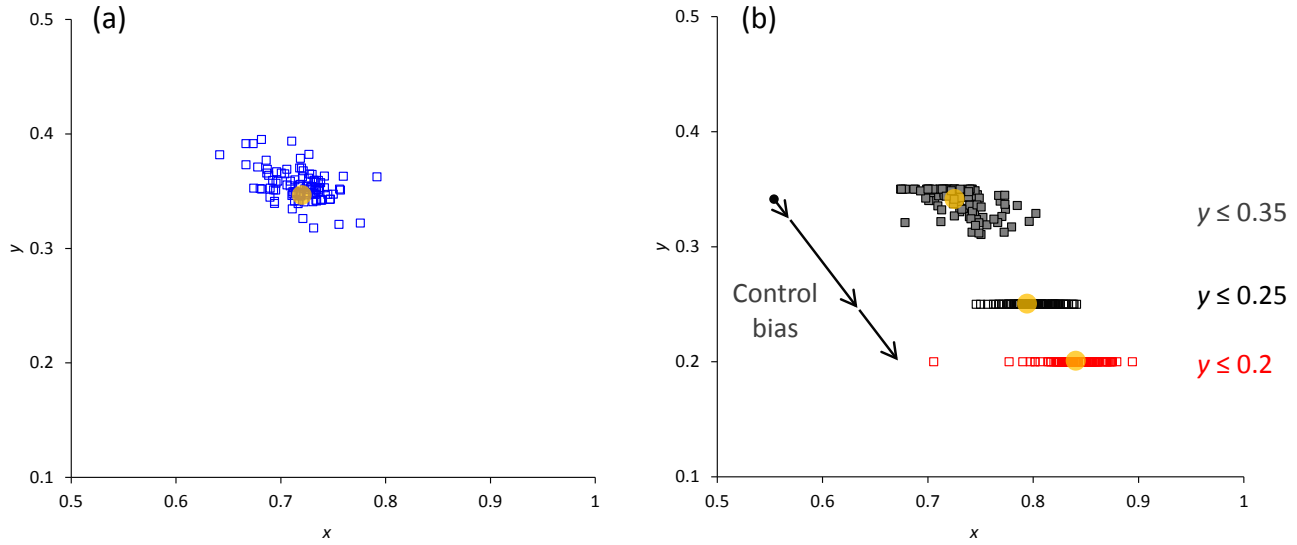


Figure 5. Numerical experiment results for the control performance of the bass-crayfish ecosystem: (a) Without limitation for the nondimensionalized biomass of bass *y*. (b) The nondimensionalized biomass of bass *y* was limited equal or lower than 0.35 (gray), 025 (black), and 0.2 (red) to mimic the decreased capacity by the climate shift. In each panel, horizontal and vertical axes mean the nondimensionalized biomass of crayfish and bass, respectively. Each rectangle means the result after *n*-th ($n \leq 100$) iterations in the calculation for the differential equations (Equation (5)). Yellow-colored circles mean the average among the 100 iterations. The control bias was defined as a vector from the average without limitation for the nondimensionalized biomass of bass *y* to the corresponding average.

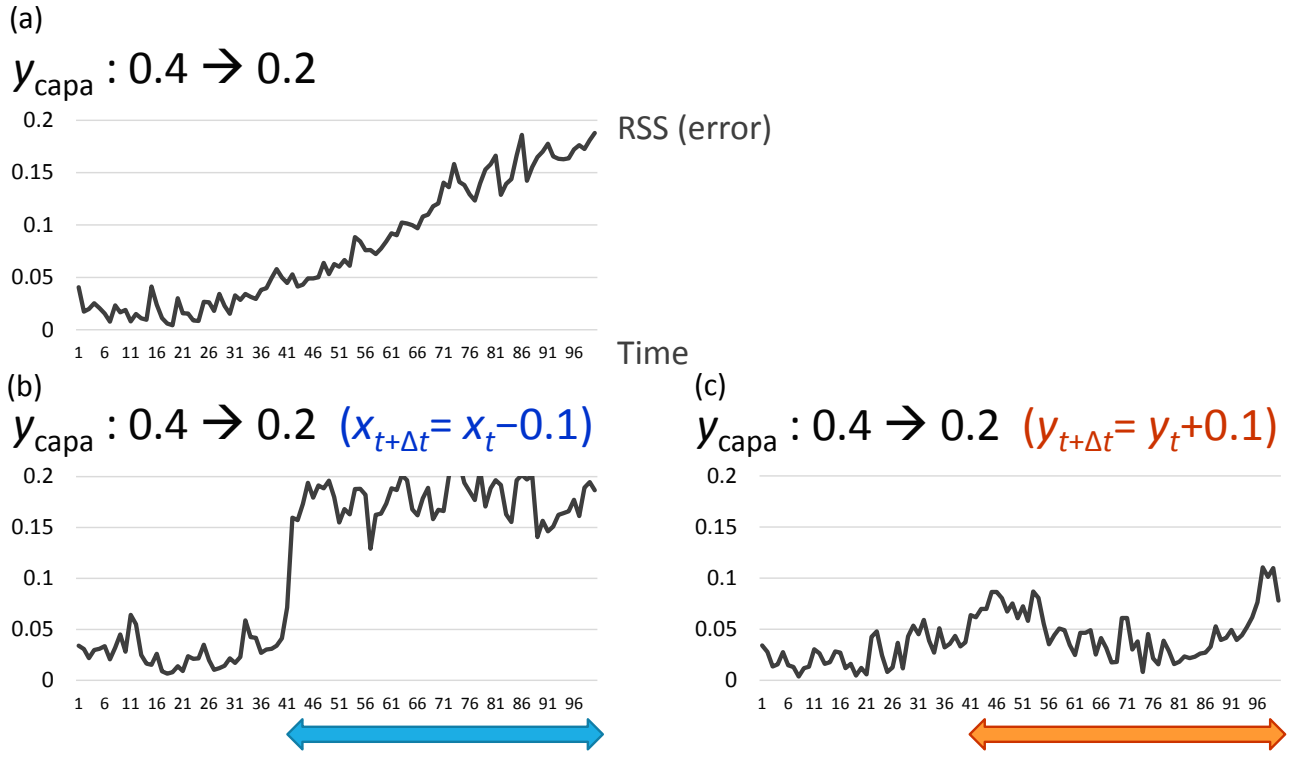


Figure 6. Numerical experiment results for the control performance of the bass-crayfish ecosystem with management actions: (a) Without the management action. (b) The nondimensionalized biomass of crayfish *x* was continuously reduced as $x_{t+\Delta t} = x_t - 0.1$ ($t>40$). (c) The nondimensionalized biomass of bass *y* was continuously increased as $y_{t+\Delta t} = y_t + 0.1$ ($t>40$), where *t* means the number of iteration in the calculation for the differential equations (Equation (5)). The horizontal and vertical axes mean the number of iteration in the calculation for the differential equations (Equation (5)) and deviation of *x* and *y* calculated by a root-sum-square of *x* and *y* from the average without limitation of the the nondimensionalized biomass of bass *y*. In (a), (b), and (c), the nondimensionalized biomass of bass *y* was limited equal or lower than 0.4 and 0.2 at $t \leq 40$ and $t>40$, respectively. The colored-arrows in (b) and (c) indicate the period in which the management action was conducted.

Novel Weight Estimation Analyses and the Development of the Wearable IngVaL System for Monitoring of Health Related Walk Parameters

Per Anders Rickard Hellstrom
Mälardalen University
Embedded Sensor Systems for Health (ESS-H)
Västerås, Sweden
e-mail: per.hellstrom@mdh.se

Mia Folke
Mälardalen University
Embedded Sensor Systems for Health (ESS-H)
Västerås, Sweden
e-mail: mia.folke@mdh.se

Abstract— **The total amount of lifted weights and lift frequency are moderate to strong risk factors for lower back pain. Measurement of carried weight is thereby of interest. The aim of this paper is to (1) present three novel analyses methods for estimation of weight during walk and (2) to describe the design process of the cost-effective research system IngVaL based on pedobarography. The paper will also (3) present the durability of the sensors. Motivations for choices in the system design are given for hardware, selection of sensor type, sensor implementation and calibration of sensors. To measure weight during walk with IngVaL, fifteen test persons made five walks each with a pseudo-random added extra weight. Three analyses methods were tested, for estimation of weight while walking, resulting in Root Mean Square Errors of 11.3 kg, 7.1 kg and 6.1 kg respectively. The durability of the sensors were tested in an outdoors walking condition. It can be concluded that the IngVaL system shows good durability and that weight during walk is possible to measure with simple analyses methods.**

Keywords- pedobarography; carried weight; portable; wearable; insole; in-shoe; personal health monitoring; measurement system design.

I. INTRODUCTION

This paper is an extension of a conference paper, which reported a Root Mean Square Error (RMSE) of 13.8 kg for monitoring of weight while walking [1].

There are numerous applications for analyses of foot plantar pressure distribution using insoles, containing pressure sensors, in the shoes. Examples are gait analysis [2], posture analysis [3], humanoid robotics [4], evaluation of footwear [5] and footwear design [6], sports [7][8], stroke rehabilitation [9][10][11] and measurements during daily human activity [12][13].

There are several types of sensors that are used in systems for foot plantar measurement [14]. Four commonly used sensor types, that are used by researchers building pedobarography systems, are the capacitive [15], the piezoelectric [16], the resistive [17][18][19] and the optoelectronic [13][19]. Three commonly used commercial portable pedobarography systems, for respective sensor type, are compared in Table 1.

Some applications for analyses of foot plantar pressure distribution need insoles with a matrix of sensors. But in some applications, when this is not needed, it is of interest to use insoles that cost less and have good durability to allow measurements in large populations over a long time. For this kind of measurements over time a pedobarography research system has been developed. First, a prototype version was designed by the authors. An improved second version was named Identifying Velocity and Load (IngVaL).

Table 1. Three commercial portable pedobarography systems compared.

System Name	Pedar™	F-Scan™	ParoTec™
Company	Novel GmbH, Munich, Germany	Tekscan, Inc., Boston, USA	Paromed GmbH, Heft, Germany
Sensor Type	Capacitive	Resistive	Optoelectronic
Number of Sensels	Up to 256	Up to 960	Up to 36
Insole Thickness	1.9 mm	0.15 mm	3.5 mm
Pressure Range	(0.03 - 1.2) MPa	Up to 0.86 MPa	Up to 0.63 MPa
Static Load Drift	Yes, compensation	Yes, no ompensation	No
Data Transfer	Bluetooth	Wi-Fi	Local storage
3-axis Sensitivity	No	No	Yes
Main Advantage	High max pressure	Fits all shoe sizes	Measures 3D forces
Main Disadvantage	Many insole sizes	Durability issue	Low max pressure

Pain in the lower back is one of the most common health problems today [20]. About a third of all employees in Sweden, during the year 2015, had pain in their lower back every week [21]. The year 2015, 16% of the employed men and 10% of the employed women in Sweden lifted more than 15 kg several times a day [21]. The total amount of lifted weights and lift frequency are moderate to strong risk factors for lower back pain [22]. Measuring of carried weight is thereby of interest. A wearable system is needed to monitor work conditions over a longer time period and preferably also during walking. Sazonova et al. have been able to measure weight using a wearable pedobarography system, but the persons had to stand still during the measurements [17].

The aim of this paper is to (1) present three novel analyses for estimation of weight during walk and (2) to describe the design process of the cost-effective pedobarography system IngVaL, especially regarding the

sensor implementation in the insoles and the dynamic calibration of the sensors. The paper will also (3) present the durability of the sensors.

II. System Development

This system development section has three subsections called Selection of Sensor Type, Sensor Implementation and Calibration of the Sensors. There are two versions of the pedobarography research system, the first prototype version and the second version, IngVaL. The design process of the two systems are described. Block diagrams of the first prototype system and the IngVaL system are shown in Figure 1.

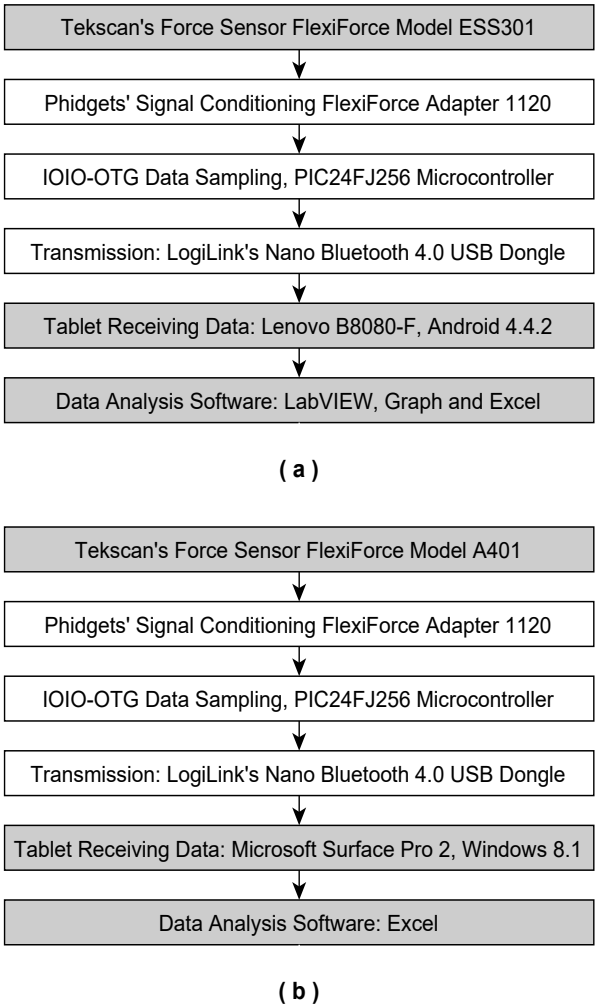


Figure 1. The two system versions; (a) the first version prototype, and (b) IngVaL, the second version of the system that was used in the experiment in this study. Grey highlights system differences.

Motivations for choices in the system design are given in the relevant hardware sub-subsections and shortcomings of the first prototype system are explained to show the evolution of the design. A comparison between the two system versions is shown in Table 2.

Table 2. Comparison between the first prototype version of the system and the IngVaL version.

System Version	Prototype (version 1)	IngVaL (version 2)
Tekscan Sensor Model	ESS301	A401
Sensor Diameter	9.5 mm	25.4 mm
Sensor Thickness	0.2 mm	0.2 mm
Insole Thickness	8 mm	6 mm
Base Insole Material	Polyurethane	Ethylene-Vinyl Acetate (EVA)
Insole Material against Foot	Polyurethane	Leather
Sensor Boundary Protection	No	Yes

Four force sensors in each insole were selected and the signal conditioning was done using the FlexiForce adapter model 1120 (Phidgets Inc., Calgary, Canada). The sampling, with a frequency of 200 Hz, and the Bluetooth transmission were done with an IOIO-OTG (Sparkfun Electronics Inc., Niwot, USA). The IOIO-OTG is based around a PIC24FJ256 microcontroller. A tablet received and stored the collected data. Software for receiving data from the IOIO-OTG is available for Android [23] and for Windows [24]. The whole system was built using commercial off the shelf components.

A. Selection of Sensor Type

Researchers have used many different types of force sensors when designing pedobarography systems and all of those sensor types have their advantages and disadvantages. Desirable sensor properties are low sensor thickness, low sensitivity for temperature change, low hysteresis, high linearity, good electrical stability and good durability.

Resistive sensors can be very thin, have good electrical stability and no hysteresis. Drawbacks are non-linearity, drift of the output when the sensor is under static load and they can have problem with durability. Piezoelectric sensors are linear, robust and not sensitive to electromagnetic interference. Main disadvantages are that it can only be used in dynamic applications, the output from the sensor is temperature dependent and it needs to have amplification close to the sensor. Capacitive sensors have a large measurement range and low temperature sensitivity but special care has to be taken to reduce parasitic capacitance when connecting the sensors to the electronics. Another disadvantage is that they have a non-linear output. Optoelectronic sensors are not sensitive to electromagnetic interference but are often thicker than the other types. They can also be used to sense forces in three dimensions but they often cannot differentiate between the three planes.

The durability of the sensors became a problem in the first prototype system. Occasional maximum readings were the first sign of sensors breaking down and thus those sensors had to be replaced and recalibrated.

The model of the force sensing resistors in the prototype version were model ESS301 (made by Tekscan Inc., Boston,

USA) with a 9.5 mm diameter of the active sensor area. The motivation for selecting this specific model was that it had good protection against humidity caused by perspiration. Larger forces can be measured when a small sensor area is used since the sensor measures the average of the forces acting on the active sensor area. Non-linearity is not a problem after calibration. Good availability and low thickness (0.2 mm) were two additional advantages. The durability of the sensors became a problem in the first prototype system.

An additional measuring problem with the first prototype system was that the sensor area was too small, which made small foot movements inside the shoe a problem. In the IngVaL system the sensor model was changed to a larger model, A401 (Tekscan Inc., Boston, USA), to reduce that problem. It also made it possible to measure a larger area of the foot. Model A401 of the sensor and the signal conditioning adapter are shown in Figure 2.

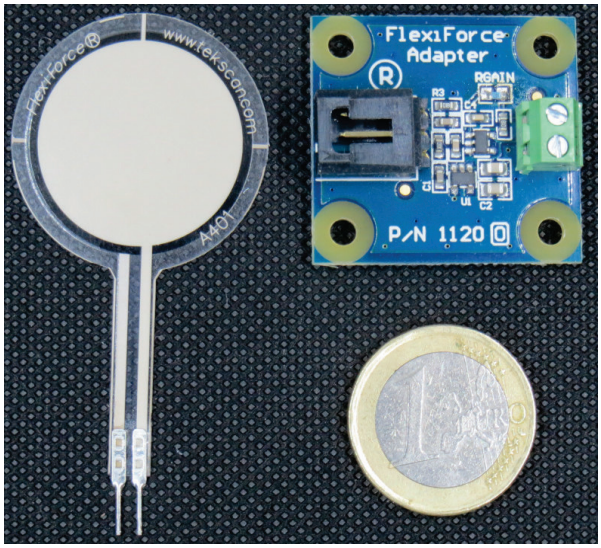


Figure 2. Sensor model A401 by Tekscan Inc., FlexiForce Adapter 1120 by Phidgets Inc. and a 1 euro coin for size comparison.

The A401 model has an active sensor diameter of 25.4 mm. The black circle on the sensor is the boundary of the active sensor area.

B. Sensor Implementation

The primary supporting positions for force interactions between the foot and the insole are the heel and the metatarsal pad (MTP), which is the pad under the forward part of the sole closest to the toes. The end of those bones, closest to the toes, are the metatarsal heads which are the bone structure for the MTP under the foot. The sensors are often placed under the heel, the MTP and the big toe, due to the bone structure of the foot and the possibility to monitor the forces from heel strike to when the foot leaves the ground again [25][26][27]. Sensors were therefore placed under the MTP and the heel, both in the first prototype system and in the IngVaL system. A fourth sensor was added to monitor when the big toe pad is in use since it is the last sensor to be activated during the stance phase (when there is contact with the ground) of the step. The four sensor locations are shown in Figure 3.

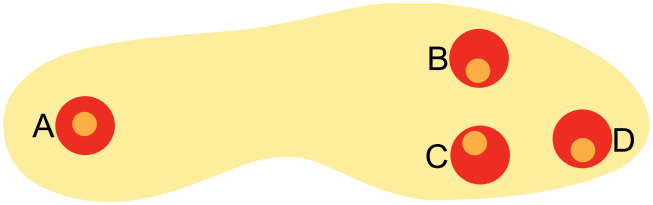


Figure 3. Sensor locations; small circles for the prototype version and big circles for IngVaL. A is the heel sensor, B is the outer and C is the inner metatarsal pad sensor, D is the big toe pad sensor.

Mechanical stress on the boundaries of the active sensor area turned out to be the culprit of the sensor breakdowns in the prototype after investigating the broken sensors. Stress on the boundaries short-circuits the sensors and results in quite rapid breakdown. The first prototype insoles were not comfortable for the user since the interface material against the foot was made of polyurethane. This material is not good at absorbing perspiration but the resulting shoe environment did not affect the sensor functionality.

New versions of the insoles were made for the second version of the system, IngVaL. The sensors were implemented sandwiched between a base of EVA and a protecting cork and leather layer of model 6949 (BNS Bergal, Nico & Solitaire, Vertriebs GmbH, Mainz, Germany). The leather interface, towards the foot, also helped reducing perspiration. Material was removed in the EVA base, under the boundaries of the active sensor areas, to remove any mechanical stress on the boundaries. The two sensors under the MTP were moved slightly more apart in the IngVaL version to better be able to measure forces close to the edges of the insoles.

C. Calibration of the Sensors

The use case for the system has to be taken into account when choosing the amplification to make sure that there is no saturation of the signal when measuring the largest forces. The amplification is changed by replacing a resistance on the Phidgets signal conditioning adapter. The manufacturer (Tekscan Inc., Boston, USA) recommends using at least four different force levels. The calibration function solves the problem with non-linearity. Sensor calibration was done by using five different force levels and repeated five times per force level. The forces were applied dynamically to avoid the problem with static load drift [28]. Polynomials of at least the third order is recommended for force sensing resistors [29]. The calibration functions for the sensors in IngVaL used fourth order polynomials, see Figure 4.

A new calibration station was designed for IngVaL. A button load cell of model CZL204E (Phidgets Inc., Calgary, Canada) was used together with the Phidgets 4-input Bridge model 1046. The calibration forces were applied, by stepping with the heel placed on the top nut, to mimic the dynamic scenario the sensors are used in during walk. Using a static load would have introduced the problem with drift over time for this type of sensor.

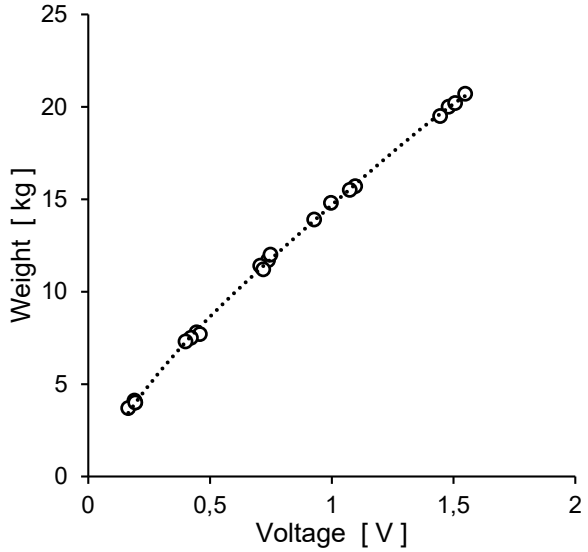


Figure 4. Weight in kilogram versus voltage in volt gives the calibration function (a polynomial of the fourth order). The calibration function for the heel's force sensing resistor is shown as an example.

A steel disk is placed over the sensing peg of the load cell to ensure a correct angle of the applied force. The angle is guaranteed because the steel disk is adapted to fit snugly on the peg of the load cell. An EVA disk is used between the heel sensor and the load cell to protect the sensor surface and to distribute the forces evenly. Metal spacers are added until the sensor almost registers pressure. Calibration forces acts perpendicular to the sensor area due to a minimized vertical displacement. The calibration station is shown in Figure 5.

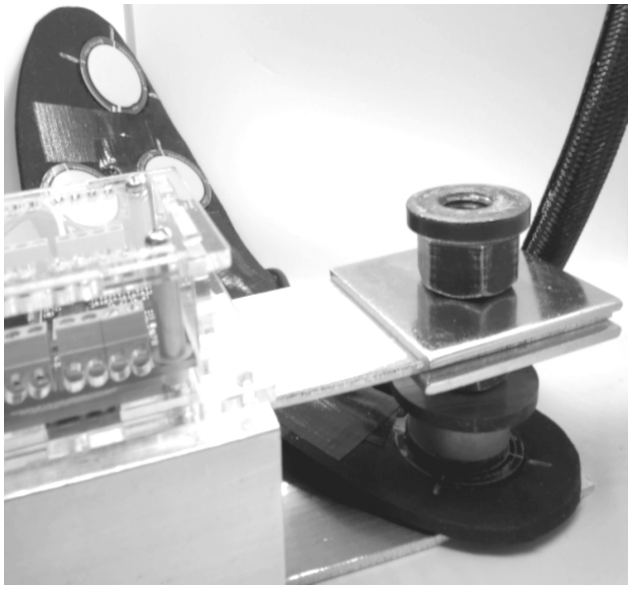


Figure 5. The calibration station used in this study, from floor; aluminium, EVA insole, force sensor, EVA disk, load cell, steel disk, upside down bolt, metal spacers, aluminium, metal spacer, nut.

The calibration function for the heel sensor is given by:

$$w(v) = -3.0588v^4 + 11.897v^3 - 17.719v^2 + 23.574v \quad (1)$$

where w is the average weight [kg] on the active sensor area and v is the voltage [V] sampled by the IOIO-OTG. The calibration terms of the functions for the sensors are shown in Table 3.

Table 3. Calibration functions, 4th order polynomials, for the force sensing resistors in the insole.

Polynomial Term	4th	3rd	2nd	1st
Heel	-3.0588	11.897	-17.719	23.574
Inside MTP	-1.9297	8.7022	-14.909	21.429
Outside MTP	-5.1492	14.801	-16.032	18.914
Big toe	-1.8116	8.4939	-15.074	21.450

III. METHOD

This section contains the experiment and data analysis. The IngVaL version of the system was used in the experiment. The experimental study has a cross-sectional design. Three novel weight estimation analyses are described in the Data Analysis subsection. Ethical approval for the study was granted by the Swedish Ethical Review Authority (diary number 2017/070).

A. Experiment

Inclusion criteria for the fifteen test persons participating in this study were to be healthy, have an EU shoe size of 43 or 44 and to be able to walk 1.0 m/s on a treadmill while carrying up to 20 kg extra weight in a backpack. The first author is shown wearing the backpack and walking on the treadmill in Figure 6.

All test persons are university staff and they had an average weight of 83.9 kg while wearing the backpack without any added weight inside it. The lowest and highest weight, without added weight in the backpack, were 75.2 kg and 110.9 kg respectively. Each test person made five walks, using the same insoles and shoes, with a pseudo-random added extra weight of (10, 20, 0, 15, 5) kg in the backpack. Safety measures were automatic stop of the treadmill (Comfort Track Prime 97690, LifeGear Ltd., Taiwan) if the test person moved away too far from the correct position, and extra padding was placed in the backpack to protect the spine. A health insurance for the test persons was bought from the Insurance department of the Legal, Financial and Administrative Services Agency. Data was recorded during one minute per walk, excluding acceleration and deceleration phases, and a test walk was made first to make sure the test person was comfortable walking on the treadmill and that all sensors were working correctly. An electronic floor scale (model GS 42 BMI, Beurer GmbH, Ulm, Germany) was used to measure the reference weight for each of the five weight configurations with an accuracy of ±0.050 kg.

Figure 6. Walk on treadmill with backpack containing the electronics of the IngVaL system and the extra carried weights.

The durability of the sensors was investigated. All sensors were in good working order during these 75 walks in the experiment. The durability of the sensors was examined further by performing additional walking outdoors.

B. Data Analysis

Three novel analyses methods for estimating weight during walk were evaluated. In analysis method 1, data from the heel sensor was used for estimation of the carried weight while walking. The 200 largest heel sensor values during each one minute walk were averaged. In analysis method 2, data from all sensors were used to estimate weight while walking. All samples from all the sensors during each one minute walk were averaged.

Fifteen test persons and five walks each resulted in 75 of these averages for each of the two analyses methods. For both methods, two averages (for 0 kg and 20 kg) for each test person were used to create a linear equation for individual calibration. The other three averages (5 kg, 10 kg and 15 kg), for each test person, were compared to the linear equation to see how big the error was. Thus, n=45 in the graph in the results section. An overview of the three steps of the data analysis are shown in Figure 7.

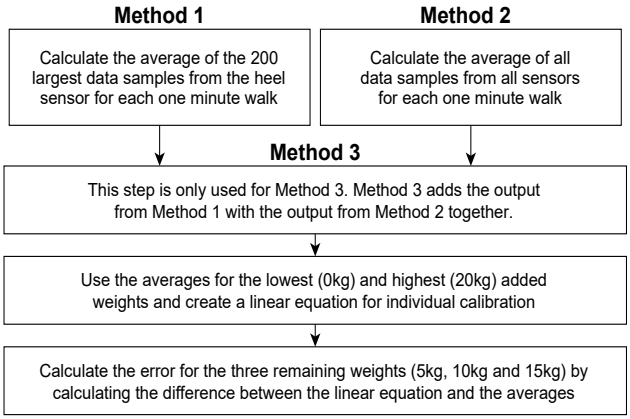


Figure 7. The steps of the data analysis, for estimation of weight during walking.

Method 1 only uses the heel sensor and the other two methods use all sensors. The outputs from method 1 and method 2 are summed together in analysis method 3.

IV. RESULTS

The first analysis method used only the heel sensor for the estimation of weight while walking, and resulted in a RMSE of 11.3 kg. The result is presented in a plot of the reference weight, as measured by the electronic floor scale, versus the weight estimation errors in Figure 8.

There is a tendency for overestimation, in method 1, of the weight with a mean of +2.1 kg. The second analysis method uses data from all the sensors and the mean is this time negative, -3.1 kg. A third method combining method 1 and method 2 seemed promising since the means have opposite signs. The third analysis method simply combines method 1 with method 2 by adding the outputs together before the individual calibration is done. The results of method 2 and method 3 are also shown in Figure 8. The result for method 3 is also shown in Figure 9 for clarity.

A summary of the results (RMSE, standard deviation (SD), mean, and upper and lower 95% confidence intervals) for the three methods are presented in Table 4.

Table 4. Summary of the results of the three methods.

Analysis Method	Method 1	Method 2	Method 3
RMSE	11.4 kg	7.1 kg	6.1 kg
n	45	45	45
SD	11.1 kg	6.5 kg	5.8 kg
Mean	2.1 kg	-3.1 kg	-2.0 kg
Upper 95% CI	23.9 kg	9.6 kg	9.3 kg
Lower 95% CI	-19.7 kg	-15.8 kg	-13.2 kg

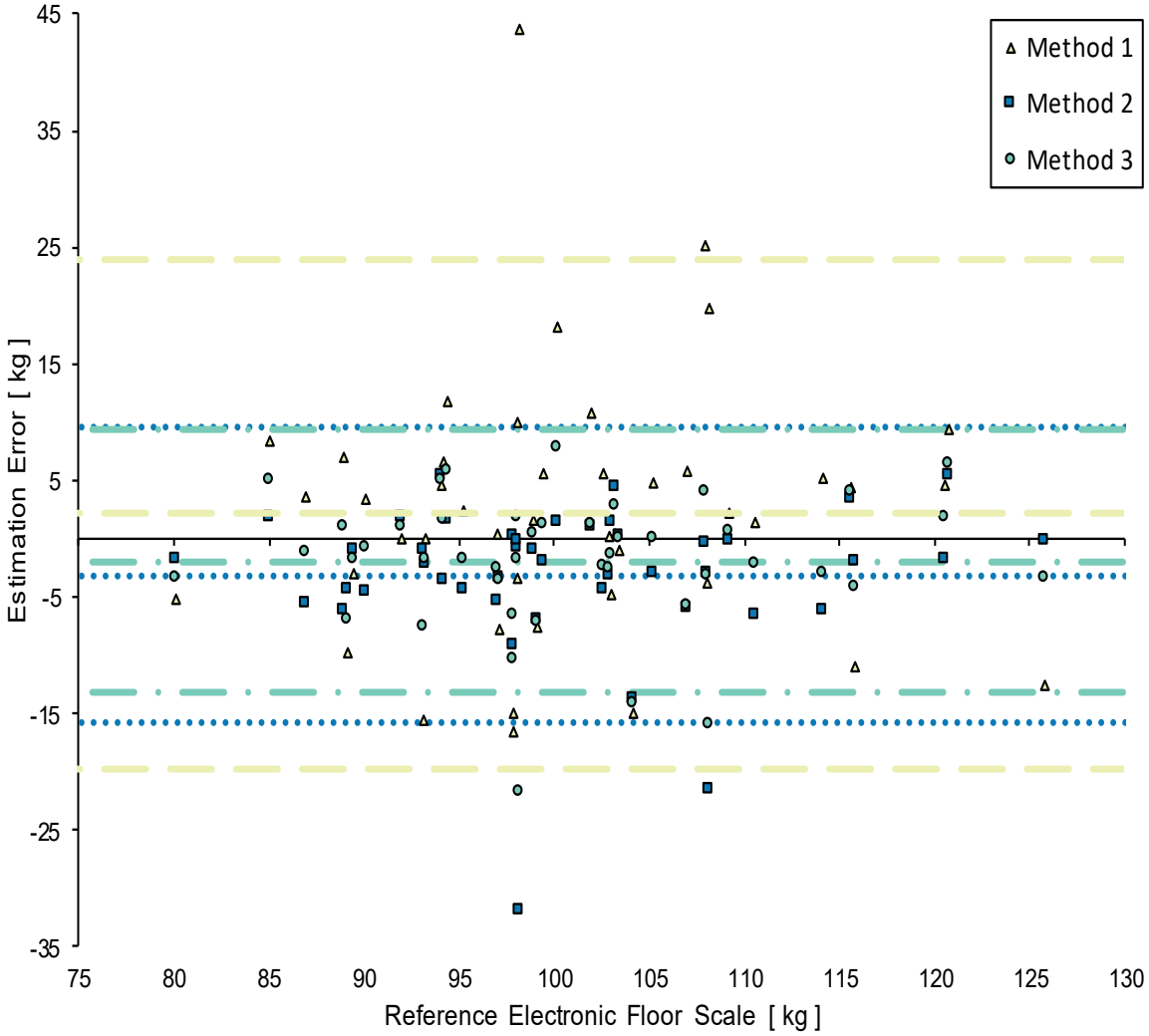


Figure 8. Errors in the estimation of the weight in kilogram versus reference electronic floor scale weight in kilogram for the three methods. Method 1 only uses heel sensor data and the other two methods uses data from all sensors.

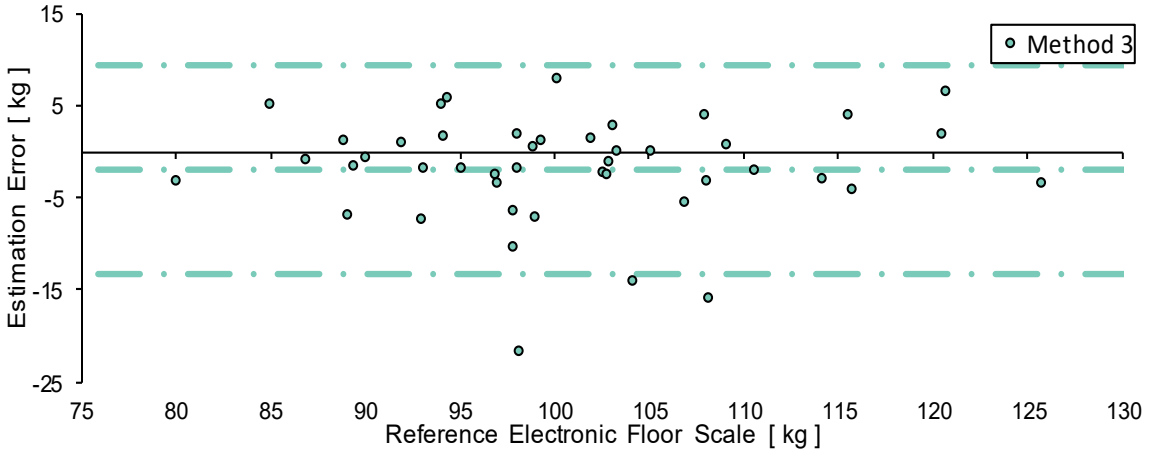


Figure 9. Errors in the estimation of the weight in kilogram versus reference electronic floor scale weight in kilogram for method 3.

Further durability testing was performed by normal walking outdoors. One sensor became unusable after a total of 36000 steps.

V. Discussion

The IngVaL system has shown to have good durability. A novel analysis method resulted in a RMSE of 6.1 kg for the weight measured during walk. The previous paper [1], which this paper is an extension of, reported a RMSE of 13.8 kg.

The earlier work was based on the equipoise method which used three sensors for the analysis and calculated a "neutral" point of balance between the forces on the heel sensor and the forward sensors [1]. One downside with the equipoise method was a very small number of usable samples for the estimation of the weight which made the estimation more uncertain.

Three analyses methods have been presented in this paper. Method 2 and method 3 used all four sensors. Method 2 resulted in a RMSE of 7.1 kg. Method 3 reduced the RMSE to 6.1 kg.

The reason for having as few sensors as possible is to reduce the cost of the system. The cost of portable commercial pedobarography systems often starts at €10000. The component cost for IngVaL would be close to €100 if the number of manufactured systems are reasonably high. Each sensor costs around €10 and this means the novel analyse method that only uses the heel sensor is a clear advantage. Method 1 only used the heel sensor and resulted in a RMSE of 11.3 kg.

The analyses methods have been kept as simple as possible. This approach was chosen in order to be able to have the analysis running on for example a smart phone or tablet in the future.

A recurring challenge, when estimation weight using few sensors, is that all persons have unique bone structures in their feet. Thus, small movements of the foot inside the shoes can result in big changes of the amount of force hitting a sensor. The approach of this paper to counter the problem of feet being unique is to do an individual calibration for each person using measured data from walk without carried weight and also from a walk with a large carried weight to create a linear equation. The larger sensor area in the IngVaL system, compared to the prototype system, also helped with this issue since a larger area of the insole is covered with sensors.

All foot arch types have proportional distribution of the forces over all regions of the foot [30]. A possible source for errors in the weight estimation is that some anatomies of feet show a more increasing contact area, between the foot and the insole, when the person carries weights compared to subjects with normal feet [31].

The most used portable pedobarography system that uses resistive sensors, F-Scan by Tekscan, states that their sensors are durable enough for multiple trials [32]. This way of expressing the durability makes it hard to compare with other systems. Great care was given in the sensor implementation in the IngVaL system to improve the durability of the sensors. The durability testing showed that the first sensor became unusable after 36000 steps. The sensor had moved a bit out of position and this put stress on the boundary of the active sensor area. Possible ways for overcoming this problem is to remove more material under the sensor, and thereby use less of the sensor area, and/or embed the sensor deeper into the insole.

Sazonova et al. estimated weight directly after coming to a standstill after walking, and reported a RMSE of 10.5 kg. Force sensing resistors drifts under static load and this adds to the challenge of measuring correctly when standing still. The drift problem is avoided when measuring during walking due to the dynamic loads on the sensors. Both methods 2 and 3 have a lower RMSE, than the system measuring during standing still.

Since the total amount of lifted weights and lift frequency are moderate to strong risk factors for lower back pain it is of interest to measure heavy working conditions over time. IngVaL can measure during walk and this means that the system is potential candidate for this monitoring. Heavy lifts are of course not the only factor for lower back pain. A sedentary lifestyle or incorrect lifting technique are examples of other factors that also can cause back problems.

VI. Conclusion and Future Work

This paper presented three novel analyses for estimation of weight during walk and described the design process of the pedobarography system. The durability of the sensors can still be improved upon. Another modification of the sensor implementation can be done to stop the sensors from moving in the insole plane and putting stress on the boundary of the active sensor area.

Acknowledgment

This research was funded by the Swedish Knowledge Foundation (KKS) through the research profile ESS-H, diary number 20120275.

Many thanks to Anders Hellström for his assistance with the manufacturing of the insoles for IngVaL and with the design of the calibration station.

References

[1] P. A. R. Hellstrom and M. Folke, "Monitoring of Carried Weight During Walk Using a Wearable Pedobarography System," The Fifth International Conference on Smart Portable, Wearable, Implantable and Disability-oriented Devices and Systems, SPWID 2019, pp. 5-8.

[2] S. Crea, M. Donati, S. M. De Rossi, C. M. Oddo, and N. Vitiello, "A wireless flexible sensorized insole for gait analysis," Sensors (Basel), vol. 14, pp. 1073-1093, 2014.

[3] E. S. Sazonov, G. Fulk, J. Hill, Y. Schutz, and R. Browning, "Monitoring of posture allocations and activities by a shoe-based wearable sensor," IEEE Trans Biomed Eng, vol. 58, pp. 983-990, 2011.

[4] A. Konno, N. Kato, S. Shirata, T. Furuta, and M. Uchiyama, "Development of a light-weight biped humanoid robot," IEEE International Conference on

Intelligent Robots and Systems, vol. 3, pp. 1565-1570, 2000.

[5] M. J. Mueller, "Application of plantar pressure assessment in footwear and insert design," J Orthop Sports Phys Ther, vol. 29, pp. 747-755, 1999.

[6] A. J. M. Boulton, C. I. Franks, R. P. Betts, T. Duckworth, and J. D. Ward, "Reduction of Abnormal Foot Pressures in Diabetic Neuropathy Using a New Polymer Insole Material," Diabetes Care, vol. 7, pp. 42-46, 1984.

[7] E. Eils et al., "Characteristic plantar pressure distribution patterns during soccer-specific movements," Am J Sports Med, vol. 32, pp. 140-145, 2004.

[8] T. Holleczek, A. Rüegg, H. Harms, and G. Tröster, "Textile pressure sensors for sports applications," SENSORS, 2010 IEEE, pp. 732-737.

[9] I. H. Khoo, P. Marayong, V. Krishnan, M. N. Balagtas, and O. Rojas, "Design of a biofeedback device for gait rehabilitation in post-stroke patients," 2015 IEEE 58th International Midwest Symposium on Circuits and Systems (MWSCAS), pp. 1-4, 2015.

[10] H. P. von Schroeder, R. D. Coutts, P. D. Lyden, E. Billings Jr., and V. L. Nickel, "Gait parameters following stroke : A practical assessment," Journal of Rehabilitation Research and Development, 1995, pp. 25-31.

[11] S. Edgar, T. Swyka, G. Fulk, and E. S. Sazonov, "Wearable shoe-based device for rehabilitation of stroke patients," Conf Proc IEEE Eng Med Biol Soc, vol. 32(1), pp. 3772-3775, 2010.

[12] M. Saito et al., "An in-shoe device to measure plantar pressure during daily human activity," Med Eng Phys, vol. 33, pp. 638-645, 2011.

[13] S. M. De Rossi et al., "Development of an in-shoe pressure-sensitive device for gait analysis," Conf Proc IEEE Eng Med Biol Soc, pp. 5637-5640, 2011.

[14] A. H. Razak, A. Zayegh, R. K. Begg, and Y. Wahab, "Foot plantar pressure measurement system: a review," Sensors (Basel), vol. 12, pp. 9884-9912, 2012.

[15] O. Mazumder, A. S. Kundu, and S. Bhaumik, "Development of wireless insole foot pressure data acquisition device," Communications, Devices and Intelligent Systems (CODIS), 2012 International Conference on, pp. 302-305, 2012.

[16] S. J. Morris and J. A. Paradiso, "Shoe-integrated sensor system for wireless gait analysis and real-time feedback," Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference. Proceedings of the Second Joint, vol. 3, pp. 2468-2469, 2002.

[17] N. A. Sazonova, R. Browning, and E. S. Sazonov, "Prediction of bodyweight and energy expenditure using point pressure and foot acceleration measurements," Open Biomed Eng J, vol. 5, pp. 110-115, 2011.

[18] S. Corbellini, "Low-cost wearable measurement system for continuous real-time pedobarography," Medical Measurements and Applications (MeMeA) 2015 IEEE International Symposium on. IEEE, pp. 639-644, 2015.

[19] M. Donati et al., "A flexible sensor technology for the distributed measurement of interaction pressure," Sensors (Basel), vol. 13, pp. 1021-1045, 2013.

[20] D. Hoy et al., "A systematic review of the global prevalence of low back pain," Arthritis Rheum, vol. 64, pp. 2028-2037, 2012.

[21] The Swedish Work Environment Authority (Arbetsmiljöverket), Arbetsmiljöstatistik 2016:2, 2016.

[22] H. Heneweer, F. Staes, G. Aufdemkampe, M. van Rijn, and L. Vanhees, "Physical activity and low back pain: a systematic review of recent literature," Eur Spine J, vol. 20, pp. 826-845, 2011.

[23] Johannes Rieke, IOIO Meter Voltage Measurement (for Android), accessed: Nov. 28, 2015, [online], available: https://github.com/jrieke/ioiometer

24] Johannes Rieke, ioiometer-pc (for Windows), accessed: Mar. 7, 2016, [online], available: https://github.com/jrieke/ioiometer-pc

[25] T. L. Lawrence and R. N. Schmidt, "Wireless in-shoe force system [for motor prosthesis]," Engineering in Medicine and Biology Society. Proceedings of the 19th Annual International Conference of the IEEE, vol. 5, pp. 2238-2241, 1997.

[26] E. S. Sazonov, N. Hegde, and W. Tang, "Development of SmartStep: an insole-based physical activity monitor," Conf Proc IEEE Eng Med Biol Soc, pp. 7209-7212, 2013.

[27] M. Chen, B. F. Huang, K. K. Lee, and Y. S. Xu, "An Intelligent Shoe-Integrated System for Plantar Pressure Measurement," Robotics and Biomimetics. ROBIO '06. IEEE International Conference on, 2006, pp. 416-421.

[28] J. Florez and A. Velasquez, "Calibration of force sensing resistors (fsr) for static and dynamic applications," 2010 IEEE ANDESCON, pp. 1-6.

[29] J. M. Brimacombe, D. R. Wilson, A. J. Hodgson, K. C. Ho, and C. Anglin, "Effect of calibration method on Tekscan sensor accuracy," J Biomech Eng, vol. 131(3), pp. 034503-034506, 2009.

[30] S. L. Goffar et al., "Changes in dynamic plantar pressure during loaded gait," Phys Ther, vol. 93, pp. 1175-1184, 2013.

[31] C. Schulze et al., "Effects of wearing different personal equipment on force distribution at the plantar surface of the foot," ScientificWorldJournal, pp. 827671-827678, 2013.

[32] Tekscan Inc. (Boston, USA), F-Scan System | Tekscan, accessed: Oct. 3, 2019, [online], available: https://www.tekscan.com/products-solutions/systems/f-scan-system

Leveraging Statistical Methods and AI Tools for Analysis of Demographic Factors of Opioid Overdose Deaths

Amna Alalawi
DMgmt in Strategic Leadership Program
Thomas Jefferson University
Philadelphia, PA, USA
amna.alalawi@jefferson.edu

Les Sztandera
Kanbar College of Design, Engineering, and Commerce
Thomas Jefferson University
Philadelphia, PA, USA
les.sztandera@jefferson.edu

***Abstract* – AI is a machine learning approach that is applied to evaluate data, makes assumptions, and facilitates a predication at different scales. It plays a key role in data science to evaluate a range of past data through statistical analysis. In similar way, big data analysis is termed as a combination of various structured, unstructured data collected by companies that would support companies to evaluate future predictions and machine learning. Deaths from drug overdose including opioid overdose have been increasing at an alarming rate, and authorities still find tackling this problem an acute challenge. This paper applies Artificial Intelligence and statistical techniques to big data to identify the demographic and socio-economic factors that have led to the increasing number of drug overdose deaths in Allegheny County, Pennsylvania, United States. Using Artificial Intelligence software, we analyzed a dataset of over 3,500 patients alongside demographic and socio-economic variables to gain detailed insights into the issue, insights that we can generalize to craft solutions to this problem in both domestic and global communities. Our findings revealed patterns ranging from possible psychological and behavioral factors and drug use on weekends, a direct market supply effect on the number of deaths, age as a factor in most overdose incidences, the majority of the population are uninsured or unemployed, and live on the outskirts of Pittsburgh. These findings imply the need for authorities to offer educational workshops to individuals and their families about the dangers of the current drug epidemic and to design an effective policy for the oversight of drug market supply that includes taking firm action against violators. The research also showed that incidences of drug overdose by age is the least among Indian, but highest among White or African American people. The study further revealed that the habit and trend of drug consumption are very common among the White and African American communities. The importance different approaches of data analytics has enhanced on significant manner to evaluate the different aspects of Drug overdose and death cases.**

Keywords – data analytics; big data; opioids; drug overdose

I. INTRODUCTION

This extended research is based on the critical global healthcare issue of examining the opioid crisis, which was presented at the Eight Data Analytics 2019 Conference in Porto, Portugal [1]. Responding to the major interests from the public discussion in Porto, this research includes more statistical analysis of the relationship between socio-demographic factors and opioid deaths in Allegheny County. Deaths caused by the opioid crisis have reached epidemic levels in the US, with more people dying from opioid overdose than by motor vehicle incidents, gun violence, or HIV [3]. In 2016, more than 42,000 Americans lost their lives as the result of a drug overdose, including 613 Allegheny County residents, a rise in deaths of over 44 percent as compared to previous year. This has largely been attributed to the presence of newer, stronger drugs such as fentanyl in US communities [10]. Local health officials have found it difficult to keep pace with the new drugs being introduced into common usage, especially the presence of fentanyl in most of the heroin sold on the streets. Fentanyl is extremely potent and can cause a fatal overdose on the first try for some users [10]. Today, drug abuse is indeed a major problem in the US. A study by the Center for Disease Control (CDC) clearly evidences the existence of this crisis, stating that in 2017, more than 47,000 people died of a drug overdose in the country [9]. It was found that more than 2,000,000 Americans live with drug addiction, such as opioid addiction [10].

This study uses Artificial Intelligence (AI) and machine learning techniques to explore a dataset containing information on 3,551 fatal incidents of opioid overdose in Allegheny County, Pennsylvania, with the aim of finding ways to minimize the consequences of the opioid crisis in communities in the US and globally.

Opioid overdose has become an epidemic in the US as well as Allegheny County [10]. In 2015, the country experienced more than 400 deaths, and from then, the trend continued to increase. Data showed that individuals affected were within the age range of 25 to 54 years old. It was found that there are plenty of opportunities that have not been taken full advantage of in terms of intervention for opioid users [17]. The authors recommended that screening for opioid and other drugs among adults involved in child welfare should be improved. Further, enhancing the ability of the direct care staff to identify opioid use and risk of overdose, as well as access to expert consultants should be increased to improve the effectiveness of the care mechanism.

In Section II of this paper, we discuss and highlight different statistical tools and methods used to carry out the investigation to study several demographic factors of those affected in Allegheny County. While Section III focuses on providing a thorough analysis of the data collected through various sources, Section IV highlights key findings of the study as well as trends in relation to the subject matter. Section V provides conclusion and recommendations pertaining to drug overdose in the community that can be used for prevention intervention. Drug overdose has continually exhibited an escalating trend in the US over the past years. From the year 1999 to 2017, an approximated number of 702,000 persons perished from drug overdose. In the year 2017 alone, about 70,000 people succumbed to death following drug overdose, making drug overdose one of the biggest causes of injury-related deaths in the US [12]. Out of these deaths, about 68% encompassed illicit opioid or prescription. The United States Department of Health and Human Services (HHS) is determined to addressing the challenges revolving around overdose, dependence and opioid and has developed five-point concrete strategy to address these issues namely; better pain management, better data, better treatment and recovery services, increased production of overdose reversers, and informed research. Based on this determination, various agencies within HHS have joined the effort. This study intends to discover how consumption of opioids and consequential deaths caused by the same are distributed amongst different demography, using Artificial Intelligence tools like data analytics and descriptive statistics.

II. METHODOLOGY

To gain greater insight into the social and economic factors that have increased the risk of fatal opioid overdose in Allegheny County, PA, we used multiple approaches utilizing AI and statistical software and programming languages, including IBM SPSS Analytics which is addressed as powerful statistical platform and it offers a great support in predictive analysis; IBM Watson Analytics that is a cloud-based smart data discovery solution that provides significant assistance in automates predictive analytics; Microsoft Power BI that is termed as a great tool to visualize a variety of data and provides more insights about different variables in data sets; Python that is identified high level general-purpose programming language; and IBM Cognos Analytics to explore the dataset, which contained information on 3,551 fatal incidents of opioid overdose in Allegheny County.

The dataset, which covers opioid overdose deaths from the year 2008 to 2017, includes fatal accidental overdoses in the county and contains information on the date of death, the time of death, the manner of death, the age, gender, and race of the decedent, the seven most prevalent drugs found in overdose victims, the zip code of the overdose incident, and the zip code of the decedent's residence [2]. To look further into this issue, we searched for other variables that may have a relationship with opioid overdose rates. These other variables can be divided into two general categories: climate and economic.

The climate variables we examined were monthly average temperature and temperature departure from mean levels for the 2008-2017 base periods. This data was retrieved using the National Oceanic and Atmospheric Administration "Climate at a Glance" tool [4]. However, there were no findings that revealed a correlation between temperature and drug overdose, thus, these graphs were eliminated from the study.

The economic indicators we examined were county unemployment rate and income inequality in the county (measured as a ratio of the mean income of the highest quintile of earners divided by the mean income of the lowest quintile of earners in the county). These two datasets were collected through the FRED Economic Data service of the St. Louis Federal Reserve Bank. The final variable examined in this category was the uninsured rate, data on which was sourced using the U.S. Census Bureau's Small Area Health Insurance Estimates program [16].

III. ANALYSIS

It has proven valuable to first observe the overdose dataset by itself to get a full understanding of the situation and to provide a baseline against which to compare the climate and economic variables examined.

The timespan of this dataset is from 01/03/08 to 12/31/17, a total of 9 years, 11 months and 28 days. The key takeaway from the data is the age range of those affected by fatal drug overdose, which is an astonishing 1 to 91. The n for this dataset is 3,460.

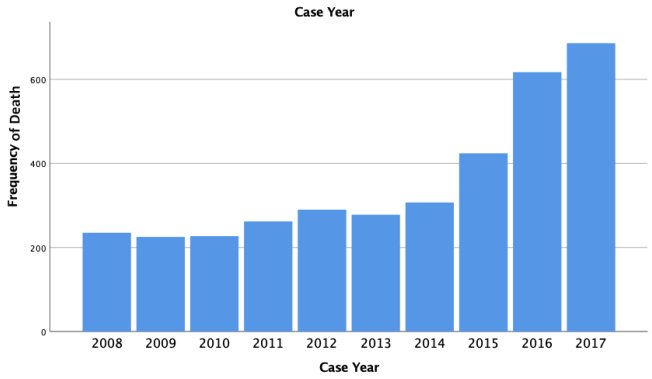


Figure 1. Overdose deaths in Allegheny Country, PA (2008-2017)

Figure 1 indicates overdose deaths over the past several years have been increasing at an exponential rate. This cannot be explained simply by an increase in population as the population of Allegheny County has fluctuated over the past decade and has, in fact, not grown at all. The graph uses data from 2008 (the earliest year in which reliable data was collected) to 2017, to illustrate the rising trend in overdose deaths.

Across all ten years covered by the dataset, it becomes clear that the crisis appears to have two distinct peaks in terms of age: one among those in their 20s and 30s and another among those in their 50s.

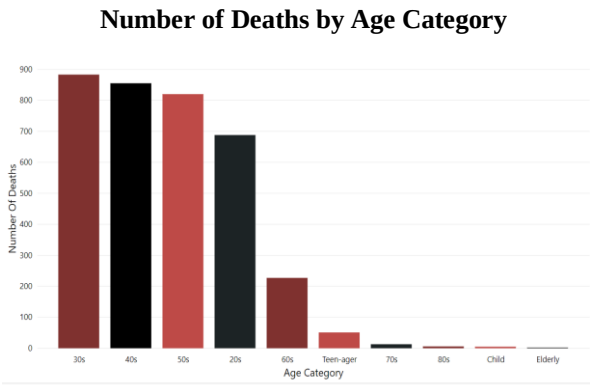


Figure 2. Number of deaths due to drug overdose by age

Figure 2 shows the number of deaths due to drug overdose by age. The most common age for someone in the county to die of an overdose is 51, indicating that there is a slight skew toward the older of the two age peaks. Moreover, the standard deviation is 12.5, showing that most deaths occur within an approximately 24-year timeframe in mid-life.

It is interesting to observe how age distribution across drug overdose deaths changes over time. The double peaks are not initially pronounced, but develop over time until 2017, when they seem to disappear. It is also interesting to see how age range widens over the years as well, and how age distribution starts to skew toward younger people. Nonetheless, the opioid epidemic is more prevalent among men than women given that women accounted for a comparably low 31 percent of fatal drug overdoses in this timeframe.

It is also clear that the overdose deaths in this county occur overwhelmingly among White Americans. However, it is worth looking at the demographics of the county overall to identify any major disparities. In Table I, we compare overdose deaths with county population data that closely aligns with U.S. Census data. While there may be some disparities in the total percentages due to non-matching categories, for the purposes of a "sanity check" on the proportions of overdose deaths in the dataset, the table serves its purpose.

It appears that Asian and Hispanic ethnic groups are comparatively less affected by overdose deaths to a small degree and African Americans are more affected, again by a small amount. Meanwhile, White American are overrepresented as victims of overdose deaths by a larger difference than any other, although still not by a significant degree.

TABLE I. PERCENTAGE OVERDOSE DEATHS BY RACE WITH COUNTY POPULATION DATA

Race	Percent of County Population	Percent of Overdose Deaths	Difference
Asian	4.0	.2	3.8%
Black or African American	13.4	14	-.06%
Hispanic	2.1	.2	1.9%
White	78.6	85.5	-6.9%

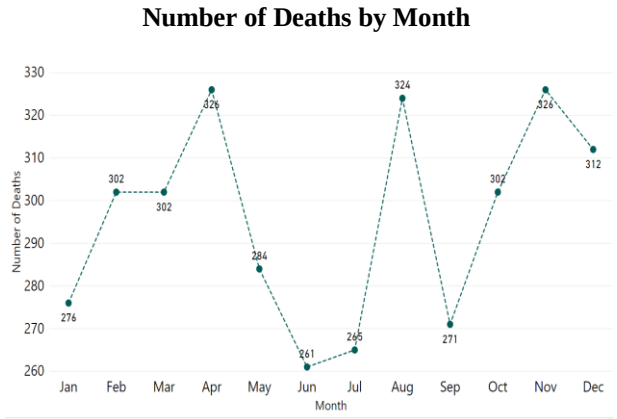


Figure 3. Number of deaths due to drug overdose by month

Figure 3 shows that April and August are the most common months when overdose deaths occur. The reasons for this are not obvious and warrant further investigation. However, other than August, there is a decline in deaths during the summer months, which may suggest a possible seasonal element in overdose deaths.

Figure 4 shows the number of deaths by day of the week, the results collected here are perhaps unsurprising. Friday, Saturday, and Sunday see the highest number of overdose deaths. People are less likely to be working on these days and will thus have more leisure time and can cause them boredom and will think of think of things to fill their time. leisure time sedentary behavior is highly associated with alcohol, tobacco, and drug use among adolescents. Saturday and Sunday are also the days when people are most social, which, depending on the person, can involve alcohol or recreational drug use.

The spike in overdose deaths that occurs at around 5 pm is also of interest and is portrayed in figure 5. This is when many people get out of work and there may be a connection here. The fact that most overdose deaths occur in the middle of the day is interesting as well, particularly the spike observed at 1 pm.

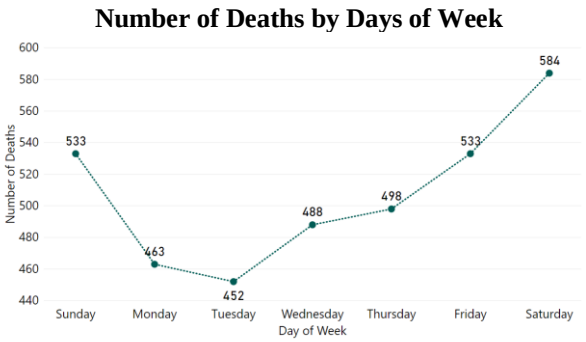


Figure 4. Number of deaths due to drug overdose by day of week

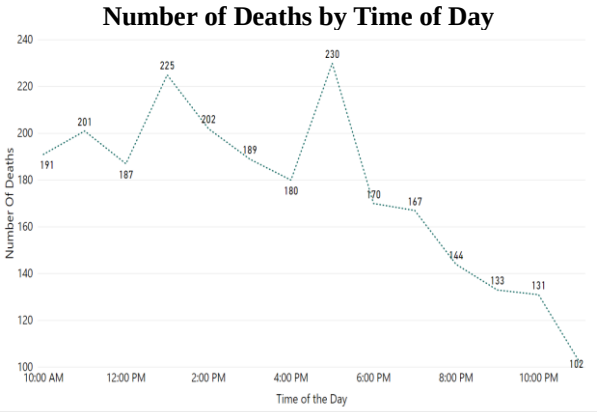


Figure 5. Number of deaths due to drug overdose by time of day

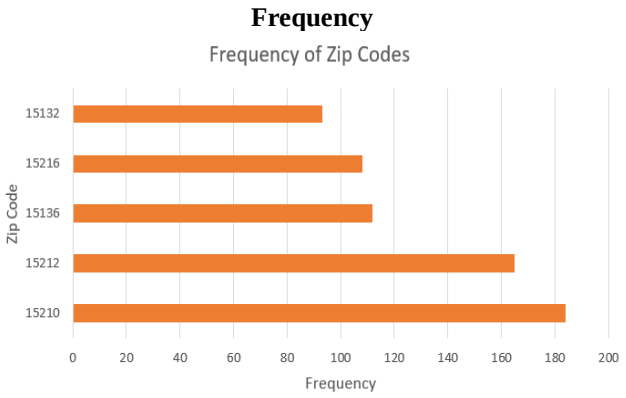


Figure 6. Top five zip codes for the locations of overdose deaths in the county during the period examined

Figure 6 shows the top five zip codes for the locations of overdose deaths in the county during the period examined. They account for a substantial proportion of all overdose deaths, but overall distribution remains quite wide and dispersed.

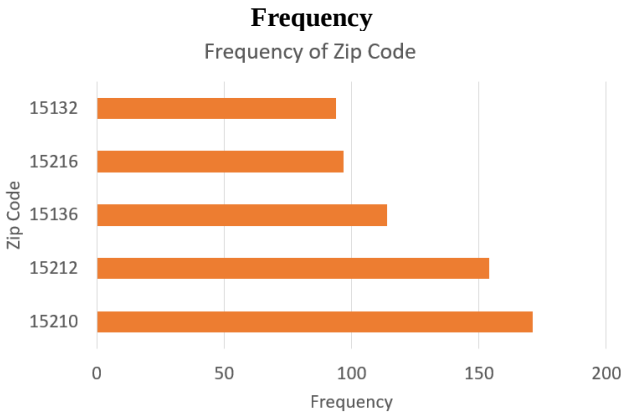


Figure 7. Top five zip codes for the residences of people who suffered a fatal drug overdose in Allegheny County during the period examined

Figure 7 shows the top five zip codes for the residences of people who suffered a fatal drug overdose in Allegheny County during the period examined. They are the same as the zip codes identified as the top five locations of fatal overdoses. This overlap indicates that people tend to overdose in the zip code in which they live, which seems reasonable. Nonetheless, there are some differences in frequency and a far larger range of zip codes cover the residences of the decedents. So, some people do travel and then overdose, including from as far away as West Virginia and even Minnesota.

Geographical Distribution of Uninsured Persons

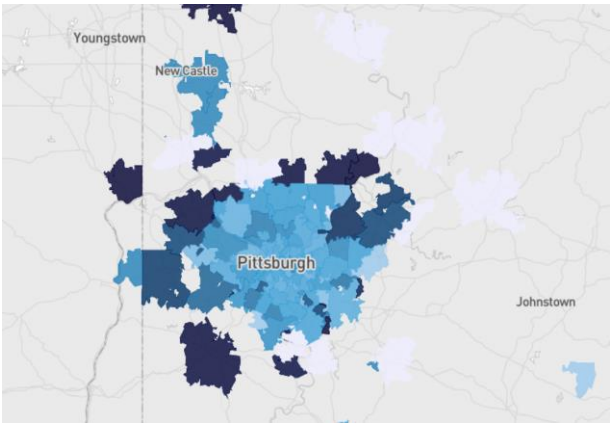


Figure 8. Map depicting the regions of the most uninsured in dark blue color verses the light blue color which shows insured people

Figure 8 shows that most uninsured people live on the outskirts of Pittsburgh. The reason behind this could be so that they can easily avoid the law and thus ensure that they

do not get caught buying and consuming drugs. In addition, since these areas are situated near the border areas, getting access to different types of opioid drugs and other such substances becomes simpler and easier for them. The dark blue color highlights the number of uninsured people. Herein it may also be said that majority of these individuals, living an uninsured life on the outskirts of Pittsburgh are also unemployed. They prefer living in this region, so that they do not have to perform any job, and so that they can gain easy access to drugs and other substances. A study released in 2017 on the lessons on the opioid epidemic states that misuse and use disorders were most common among those who were uninsured or unemployed, were low-income individuals, or had behavioral health problems [8].

IV. FINDINGS AND TRENDS

This section highlights key findings of the study, as well as trends in relation to the subject matter as per the demographic variables tested.

A. Gender and Race in Allegheny County and Overdose Deaths

Drug Overdose Deaths by Gender

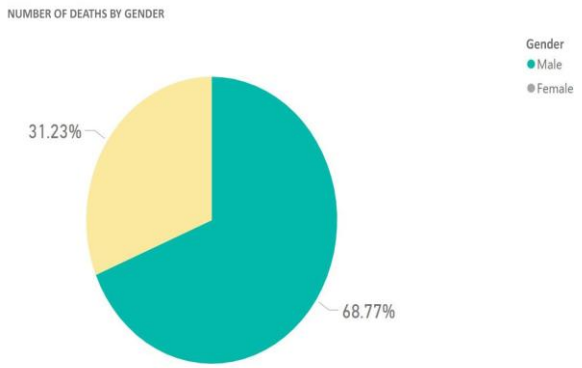


Figure 9. A chart showing the number of deaths by gender in Allegheny County, PA

Figure 9 analyzes the distribution between male and female in overdose deaths in Allegheny County. The study revalidates that overdose deaths are more amongst men, which make up around 69 percent of the total deaths. Women account for 31 percent of the total deaths. One can argue that the likelihood of good health or the prevalence of certain disorders is, in part, a product of gender. Certain health issues are unique to men; others affect men disproportionately compared to women; and still others have a different effect on men than on women. Examining gender differences is important in a study like this to help researchers in the future determine the overall co-factors that cause the development of drug overdose, and then death, and can shed the light on new insights.

Drug Overdose Deaths by Race

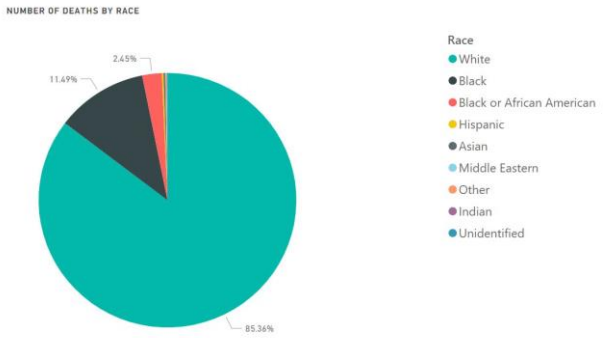


Figure 10. A pie chart showing number of deaths by race in Allegheny County, PA

Figure 10 identifies the number of deaths by race. It is seen that White Americans represent the highest segment of the population that is affected among drug users and number of fatalities due to the opioid crisis. Racial groups usually have unique features that further influence their engagement in overdose drugs. This analysis can also give a general racial background of the county of Allegheny in Pennsylvania which is known for its high White American population, and this can infer on the socio-economic and demographic status of the county, that it is mostly populated by White Americans. This analysis can create a limitation to the research.

Drug Overdose by Age & Race Distribution

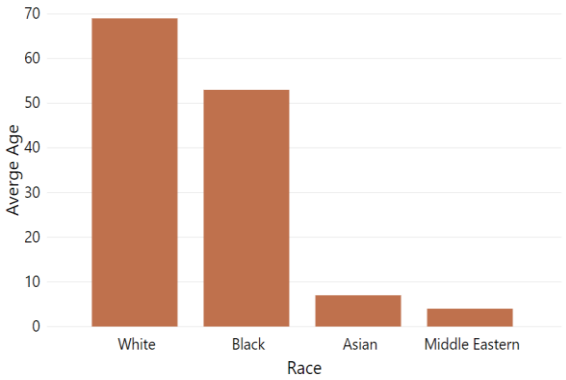


Figure 11. A bar graph showing drug overdose by age and race distribution in Allegheny County on four races

Figure 11 shows that incidences of drug overdose by age is the least among Asians, but highest among White American and African American people. The average age of these individuals is 45 years. On this basis, it can be said that 45 years old White or Black Americans who live in Allegheny County are high-risk individuals for drug

overdose, and even though Allegheny County is a developed region, there are a considerable number of people who still are taking excessive amounts of drugs. Similarly, the average age of Middle Eastern people is also 40 years. Cases of drug overdose among the Middle Eastern people is considerably high, and it is also considered one of the major markets for drug consumption. From the above information it can also be observed that these individuals start taking drugs from a very young age.

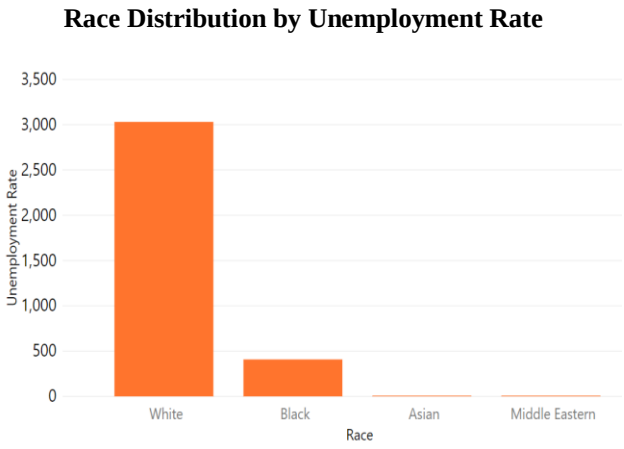


Figure 12. A bar graph showing drug overdose by unemployment rate

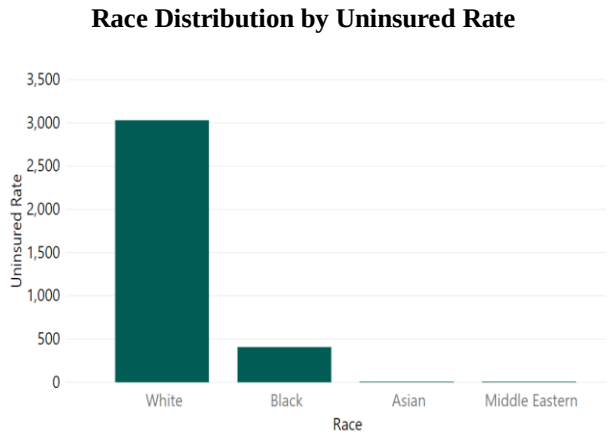


Figure 13. A bar graph showing drug overdose by uninsured rate

Figures 12 and 13 show how the drug overdose death rate is influenced by the uninsured rate and unemployment. The unemployment rate determines the percentage of total labor force that is unemployed but actively seeking employment. From the above images, it can be observed that the uninsured and unemployment rates are considerably high among White Americans. This resulted in a significant increase in the death rate. At the highest level of uninsured rate, i.e. 10%, the number of deaths was very low. However, at the lowest uninsured rate, the rate of unemployment is very high. On this basis, it can be said that there is a direct relationship between uninsured and unemployment rate. Unemployed individuals are not able to afford to insure themselves. This creates a significant gap in the number of people who have an insurance cover and the ones who do not.

A. B. Month to Month Fluctuations

Looking at every year separately on IBM Watson, a cloud system for data analytics, another pattern emerges whereby months with abnormally high death counts are followed by months with abnormally low death counts, and vice versa. This may be related to the availability of drugs in the market so as the high level of addiction to drugs related to fentanyl and heroin has identified. The authorities have successfully taken down various platforms that were previously used to sell and buy drugs, but not a lot of information has been gathered in terms of the effects of such an action [5].

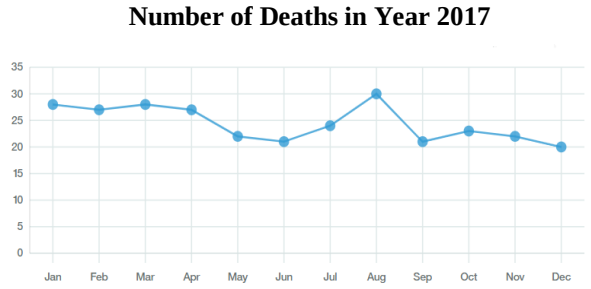


Figure 14. A visualization of the death rates of each month of the year 2017

In 2017, the aforementioned pattern was most visible in the months of August and September. August saw a high number of deaths and was followed by September, a month with lower deaths, which is the pattern we recognized through IBM Watson which we concluded that this could be related to market supply. When one particular month has a high death number, the following month is much lower, suggesting that the supply is lower in the market. As per the above chart, the month of August has higher number of death rate as compared to September. The year ended with relatively low death rates in comparison with previous months. As indicated above, this pattern was seen most visible in the months of August and September, even though we can see similar patterns in January and February, and then April and May. The conclusion is that drug addiction is so high that when drugs are available in the market, they are all consumed, which causes a shorter supply the following month, where drug overdose deaths slightly decrease.

Number of Deaths in Year 2013

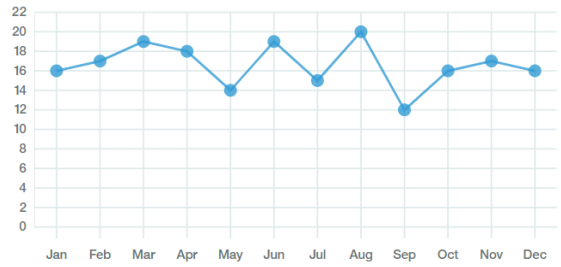


Figure 15. A visualization of the death rates of each month of the year 2013

As per the visualization in Figure 15, the year 2013 is a very good example of the potential oscillating trend, in which months with very high fatal drug overdose rates are followed by months with significantly lower overdose death rates.

Number of Deaths by Month (2008-2017)

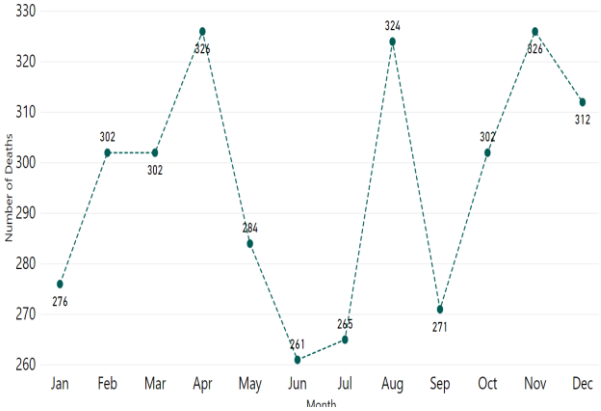


Figure 16. A visualization of the number of deaths for each month of all the years of observation

Looking at all the years examined together in Figure 16, we can confirm the pattern that shows that it may be due to the supply of drugs in the market, which causes fluctuation in deaths by months. Looking at August, for example, we see a peak in deaths, of which then September shows much lower deaths. As previously mentioned, this was also apparent when we studied each year separately on IBM Watson. There was a high indication that after a month of high deaths, the following month sees much lower deaths. The visualization in Figure 16 was generated using Microsoft Power BI.

Consumption of Drug Category OD1 by Race

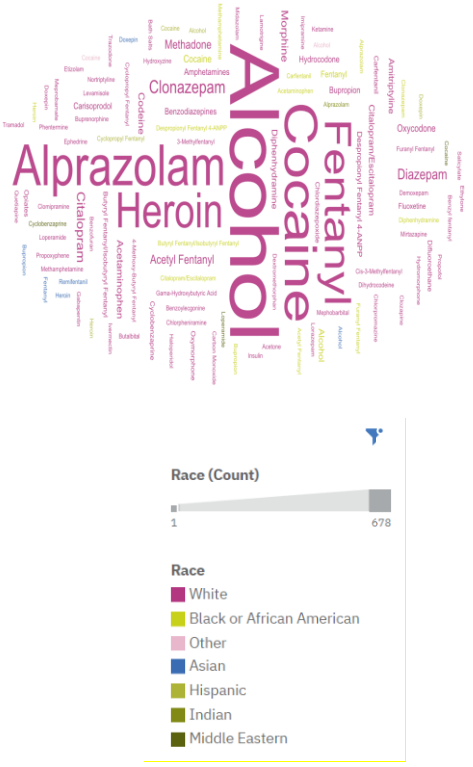


Figure 17. A word cloud of the drug names consumed in the first category which was labeled as OD1.

Figure 17 shows the consumption of one of the drug categories which was tested as part of the data analysis, and labeled as OD1, OD which stands for overdose, and had listed the names of the drugs which the patient consumed. We tested the drug names by race consumption. In the above image, which is a word cloud generated from IBM Cognos Analytics, it can be noted that White American people consume the most amounts of drugs. A word cloud is a technology to visualize language or text data, which has recently gained increasing attention and more application opportunities in big data analysis [7]. It includes the likes of Fentanyl, Heroin, Alprazolam, Alcohol, Cocaine, and many more. A recent study confirms that Fentanyl-related overdose is on the rise nationwide and is a particularly pressing problem in the eastern part of the United States. The eastern U.S. heroin market consists largely of white powder heroin, which is easier to cut with fentanyl due to its color and consistency compared to western black tar heroin. In addition, Fentanyl is easier to produce than heroin, and can be smuggled in small batches, and offers high profit margins for sellers. Thus, these characteristics have contributed to its increased prevalence in U.S. drug markets [12]. Consuming drug has almost become a common practice among White people. Due to this reason, it may not be wrong to say that these individuals are the most exposed

to different health-related issues and deaths due to drug addiction and abuse. Consuming drug is also fairly common among people from Black or African American race. However, on the other hand, Asian people consume the least amount of OD1 drugs, as seen in the figure above. Similarly, consuming OD1 drugs is almost non-existent in Indian, and Middle Eastern people. People from other races that were not included in the current study also consumed the OD1 drug in a significant number.

Consumption of Drug Category OD2 by Race

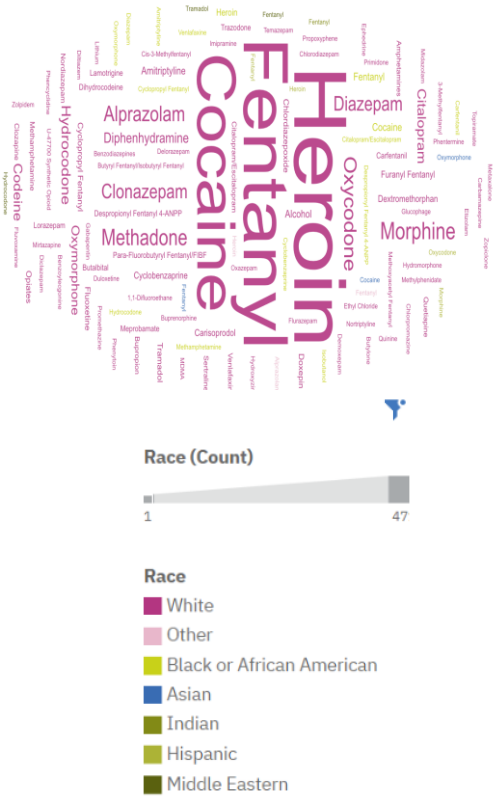


Figure 18. A word cloud of the drug names consumed in the first category which was labeled as OD2

Figure 18 shows the consumption of OD2, another category which was listed and labeled for drug names, and has been leveraged through IBM Cognos Analytics to analyze drug name by race, to reveal whether there is a pattern that is recognized. Herein it can be observed that although White Americans do not consume OD2 drugs as they consume OD1 drugs, the drugs they intake are considered among the most lethal drugs. Due to this reason, they are most exposed to deaths and other health issues from OD2 drugs. The pattern of consuming drugs such as Heroin, Cocaine, and others among White American people is increasing at a very fast pace. There are more individuals consuming such drugs than people from any other of the races. However, similar to the consumption pattern for OD1 drugs, individuals from other races, not included in the current study, also consume a wide variety of OD2 drugs, which puts their life at great risks.

Moreover, more people from Black or African American community consume OD2 drugs as opposed to OD1 drugs. On this basis, it can be observed that different races have developed different preferences for the drugs that they aim at consuming. This is weakening these communities and is putting their lives in danger. Consumption of OD2 category drugs in Asian, Indian and Middle Eastern people is again very low.

V. Conclusion and Recommendations

This data analytics study provides an expanded exploration of the problem of opioid drug overdoses in Allegheny County. Applying several statistical techniques, including pattern recognition, and generating other data visualizations, we were able to validate previously identified findings about overdose deaths, such as the age range and general demographics of affected populations [6][17]. We found the tools we used to be very useful in helping us to gain a better understanding of our sample set and to generate informative and understandable visuals.

Data Analytics also allowed us to find new information about our sample, particularly when combined with the additional variables added to the base dataset. We were able to generate highly practical visuals and illustrate clear trends from a large dataset with many variables, indicating that there is a possible relationship between drug addiction and lower temperatures and between psychological and behavioral factors and weekend drug use as well as a direct market supply effect on the number of deaths. These findings imply the need for authorities to offer educational workshops to individuals and their families, as well as health practitioners, about the dangers of the current drug epidemic and to design an effective policy for the oversight of drug market supply that includes taking firm action against violators.

The above analysis can be concluded that majority of the people, who are uninsured and unemployed, live on the outskirts of Pittsburgh so that they can get easy access to drugs. The study further revealed that the habit and trend of drug consumption are very common among Black or African American people. The average age of such individuals is 45 years meaning that they continue to consume drugs for a considerable period of their lives. During the study it was further noted that that uninsured rate has a direct relation with the rate of unemployment. This means that most of the unemployed people tend to avoid taking an insurance cover for the fear of paying money that they are not earning. In addition, they know that drug consumption will result in them being unable to take full advantage of their insurance cover if they had one.

We recommend that the authorities spend more time and funds on advertisements to educate individuals and families about the problem as well as investing in

approaches to oversee market supply to drug users. Creation of awareness is one of the most comprehensive approaches to sensitize communities concerning the consequences of opioid overdoses, but also one of the avenues to equip dependents, addicts and the general public with knowledge, skills and correct attitudes towards opioid-free and dependence-free lifestyles. The relevant governmental departments such as the Human and Health Services and other relevant institutions should vest time and resources in the creation of awareness to induce more information to the general public about the negative impacts associated with drug dependencies and also provide necessary information which can guarantee abstinence from opioid overdoses. The awareness will extend further to create an understanding about how to efficiently deal with opioid use disorder and how to deal with it effectively and comprehensively.

Besides creation of awareness, this study also recommends community empowerment programs to keep the vulnerable groups (especially the youths) occupied and attached to meaningful economic and social constructs, to help them deviate away from indulgence in opioid reliance and from engaging in other malpractices. Such empowerment programs should be tailored to alleviate laxity amongst the youths in an urbanized setting, to deviate their attention ways from any engagements in substance abuse. By so doing, this will yield more multiplier effects such as employment creation, reduced deaths and increased productivity within the existing population.

The third and last recommendation is to consider sending the victims of opioid dependence to rehabilitation centers to permanently ensure individuals undergoing detoxification do not relapse back into opioids. Opioids may rewire the brain, inducing forth substantial cravings which may trigger irresistible urges to fall back and use these drugs. The role of the rehabilitation centers will be geared to ensure persons are keenly guided and monitored towards permanent emancipation from the wants and chains of opioids and related effects.

The issue of drug overdose deaths is so pressing that more research by data and population scientists is needed to gain further insight into this epidemic, so that impactful solutions can be found to reduce its harmful effects on communities both in the United States and around the globe. This study cognizes a limited study in the field of drug overdose, especially pertaining to the causes leading to overdoses. There is need for a more systematic review and primary studies unraveling this important field which is otherwise claiming the lives of the most productive ages in a population. More study in this field will provide useful insights to policy makers, planners and the government on the way forward.

Results of the current study supported results of the past studies. Thus, the limitation of this research was that there were minimal new findings obtained. The tools used in this study helped in enabling us to recognize patterns through data visualizations, and we encourage future researchers to leverage the use of multi-dimensional data to find factors that could possibly correlate to the drug addiction epidemic to explore new findings within regional and global communities. It is the hope that this information will be used in enhancing understanding of readers, policy makers, health professionals regarding the subject matter.

REFERENCES

[1] A. Alalawi, D. Fooks, L. Sztandera and S. Zakrzewski, "Leveraging Statistical Methods for an Analysis of Demographic Factors of Opioid Overdose Deaths," The Eighth International Conference on Data Analytics, IARIA, Proceedings of 2019. pp. 49-52.

[2] "Allegheny County Fatal Accidental Overdoses," 2018. [Online]. URL: https://catalog.data.gov/dataset/allegheny-county-fatal-accidental-overdoses [accessed: 2019-04-01].

[3] D. Ciccarone, "Fentanyl in the US Heroin Supply: A Rapidly Changing Risk Environment", The International Journal on Drug Policy, vol. 46, pp. 107-111, 2017.

[4] "Climate at a Glance," [Online]. URL: http://www.ncdc.noaa.gov/cag/county/time-series/PA-003/tavg/all/1/20082017?base_prd=true&firstbaseyear=2008&lastbaseyear=2017&trend=true&trend_base=10&firsttrendyear=1895&lasttrendyear=2019 [accessed: 2019-04-01].

[5] B. Edlin, "Access to Treatment for Hepatitis C Virus Infection: Time to Put Patients First", The Lancet Infectious Diseases, vol. 16(9). doi:10.1016/s1473-3099(16)30005-6, 2016.

[6] "Income Inequality in Allegheny County, PA". 2018. [Online]. URL: https://fred.stlouisfed.org/series/2020RATIO042003 [accessed: 2019-04-01].

[7] Jin, Y. "Development of Word Cloud Generator Software Based on Python," *Procedia Engineering*, 174, 788-792, 2017.

[8] H. Koh, "Community-Based Prevention and Strategies for the Opioid Crisis," 2017. [Online]. URL: https://jamanetwork.com/journals/jama/article-abstract/2654370 [accessed: 2019-12-01].

[9] G. Lopez, "A new study shows America's drug overdose crisis is by far the worst among wealthy countries," 2019. [Online]. URL: https://www.vox.com/science-and-health/2019/2/26/18234863/drug-overdose-death-america-international-study [accessed: 2019-04-01].

[10] R. Lord. "Allegheny County drug overdose deaths surge to 613 in 2016, breaking record," 2017. [Online]. URL: http://www.post-gazette.com/news/overdosed/2017/04/06/Allegheny-County-drug-overdoses-610-deaths-break-record/stories/201704060199. [accessed: 2019-05-01].

[11] "Percent Without Health Insurance Data for Allegheny County," [Online]. URL: http://www.opendatanetwork.com/entity/0500000US42003/Allegheny_County_PA/health.health_insurance.pctui?year=2014&age=18 to 64&race=All races & gender=Both genderes & income=All income levels [accessed: 2019-05-01].

[12] B. Rhodes, B. Costenbader, L. Wilson, R. Hershow, J. Carroll, W. Zule, and L. Brinkley-Rubinstein, "Urban, individuals of color are impacted by fentanyl-contaminated heroin". International Journal of Drug Policy, 73, 1-6, 2019.

[13] B. Saloner, "A public health strategy for the opioid crisis," Public Health Reports, vol. 133, pp. 24S-34S, 2018.

[14] L. Sederer, "Take Action Against Addiction," [Online]. URL: https://www.usnews.com/opinion/blogs/policy-dose/articles/2016-02-01/10-ways-to-combat-americas-drug-abuse-problem [accessed: 2019-05-01].

[15] "Unemployment Rate in Allegheny County, PA," 2019. [Online]. URL: https://fred.stlouisfed.org/seri [accessed: 2019-05-05].

[16] "U.S. Census Bureau's Small Area Health Insurance Estimates Program," [Online]. URL: https://www.opendatanetwork.com/entity/0500000US42003/Allegheny_County_PA/health.health_insurance.nui?ref=entityquestion&year=2014&age=18%20to%2064&race=All%20races&gender=Both%20genderes&income=All%20income%20levels [accessed: 2019-04-01].

[17] E. Hulsey, and et. al., "Opiate-Related Overdose Deaths in Allegheny County – Risks and Opportunities," 2019. [Online]. URL: https://alleghenycounty.us/uploadedFiles/Allegheny_Home/Health_Department/Programs/Special_Initiatives/Overdose_Prevention/Opiate-Related_Overdose_Deaths_in_Allegheny_County.pdf [accessed: 2019-08-01].

International Patient Summary Standard Based on Archetype Concepts

Evgeniy Krastev
Faculty of Mathematics and Informatics
Sofia University St. Kliment Ohridsky
Sofia, Bulgaria
e-mail: eck@fmi.uni-sofia.bg

Dimitar Tcharaktchiev
Department of Medical Informatics
Medical University
Sofia, Bulgaria
e-mail: dimitardt@gmail.com

Petko Kovatchev
Faculty of Mathematics and Informatics
Sofia University St. Kliment Ohridsky
Sofia, Bulgaria
e-mail: az@petko.info

Simeon Abanos
Faculty of Mathematics and Informatics
Sofia University St. Kliment Ohridsky
Sofia, Bulgaria
e-mail: simeonabanos@gmail.com

***Abstract*—The design and software implementation of a standard for International Patient Summary (IPS) is recently in the focus of several updates of major eHealth standards and technical specifications in European Union countries. The design goal of the final draft version of this European standard is to be implementation independent and to enable semantic interoperability in exchanging clinical data. The paper makes a detailed analysis of the relations of this standard with other European projects. The objective of this paper is to implement the IPS standard making use of archetype concepts and demonstrates the implementation in a fully functional web application for exchange of IPS clinical data. The paper makes a detailed review of the dataset structure and the patterns used to describe the sections in the IPS standard focusing on the Medication summary section. On this basis two archetype models are created correspondingly with EN ISO 13606 and openEHR archetypes. A client-server web application is developed to demonstrate the practical application of the archetype model using openEHR specifications. The discussion of the computer experiments leads to the conclusion that archetype models of the IPS standard can fully satisfy the objectives of this standard for cross- border semantic interoperability. The obtained results correspond to existing sample implementations of the IPS with Message paradigm technologies. These results are novel because for the first time they demonstrate the implementation in a client- server application of the IPS standard designed with archetypes.**

Keywords-semantic interoperability; eHealth; international patient summary; medication summary; archetype concept.

I. INTRODUCTION

Today people, businesses and data resources are globally connected. Every day a lot of people travel around the world or simply move from place to another in the country of their origin. This active way of life of modern people relates to the need of exchanging at least a minimum set of patient data between healthcare professionals. The availability of such data at the right time in case of unscheduled and in scheduled care is essential for providing quality health services. For example, in the outbreak of dangerous virus in one country, a person leaving this country may suddenly fell ill in another country. Then the assignment of proper treatment for this patient requires interoperability of services allowing to obtain among the rest the history of illnesses, allergies and the medications he or she used to take in the country of permanent stay. The requirement for establishing and maintaining a standard International Patient Summary (IPS) becomes imperative in times when huge amounts of people migrate from country, seek employment or better life abroad.

The research work for the development of a standard for an International Patient Summary (IPS) has a long history [1]. It has started with the European Patients – Smart Open Services (epSOS) project having as a main objective to enable a service infrastructure for cross-border interoperability between Electronic Health Records (EHR) systems in Europe [2] [3]. This pilot infrastructure has been planned to allow a citizen from one EU country to receive relevant treatment for unscheduled health need in another country. The outcomes of this project have laid the foundations for sharing and exchanging patient summary and electronic prescription records. Research work has continued in several other EU projects like the Joint Action to support the eHealth Network project (JAseHN [4]) and the obtained results have been implemented in the eHealth Digital Service Infrastructure (eHDSI [5]) or adopted by the eHealth Network (eHN [6] [7]).

The European Commission for Standardization (CEN) in a collaboration with HL7 [8] [9] produced final draft versions correspondingly of a standard (EN 17269 [10]) and a technical specification (FprCEN/TS 17288 [11]) for an IPS. These two documents provide a detailed abstract specification of an IPS model from which concrete models can be derived and implemented. Therefore, the IPS model is described in terms of clinically relevant data set that is "*minimal*", "*specialty-agnostic and condition- independent*".

The objective of this paper is to investigate these newly published documents from the perspective for software implementation of the proposed domain information model and the dataset specifications. For example, it is important to learn how well this IPS model can be expressed with EN ISO 13606 and openEHR [12]., where EN ISO 13606 [13] is the EU approved standard for semantic interoperability. The reason to explore this subject in EN 17269 is that the collaboration of CEN with HL7 has led to a rather unbalanced interpretation of the use cases in terms of software technologies exclusively related to HL7 FHIR [8] and HL7 CDA [14]. Indeed the data set references to ISO 21090 [15]

make this model compatible with the data sets used in information models based on American Standards HL7 v3, HL7 FHIR or an EU standard like EN ISO 13606. On the other side, the IPS draft standard prEN 17269 does not provide guidance for implementing semantic interoperability in the exchange of IPS records. In fact, the draft version of FprCEN/TS 17288 provides just a short informative reference to EN ISO 13606 and openEHR archetypes, while semantic interoperability appears to be out of the scope of this IPS model. Moreover, the evaluation of the IPS model is presented only in terms of the Messages Paradigm of HL7 CDA and HL7 FHIR making use of a specific ART-DECOR template exchange format with Native XML databases (NXD) [16]. Besides, practical experience shows that the support for semantic interoperability of the Messages paradigm is problematic and it is difficult to scale it at national levels [17].

Our approach to evaluate the IPS model is based on designing and implementing the IPS document with archetype concepts satisfying the Archetype Object Models of EN ISO 13606 [18] [19] and openEHR [12]. It is a novel research work because it aims to evaluate newly published draft versions of a standard and the accompanying it technical specification for IPS, where the Archetype paradigm is superficially taken in consideration. Note that the ART-DÉCOR template format is not directly compatible with the Archetype Description Language (ADL) [19] [12]. The above formulated issues are currently poorly explored in the existing literature especially regarding their software implementation. In this paper, we take one of the required sections of the IPS as a case study to validate the application of our approach for introducing semantic interoperability support in the scenarios for managing IPS extracts. The details of our approach are as follows:

- Implement the IPS Medication section in EN 17269 both in terms of EN ISO 13606 Archetype Object Model (AOM)and in terms of the openEHR AOM.
- Explore the compatibility of the obtained EN ISO 13606 Archetype conceptual design with respect to the requirements of an openEHR engine for running openEHR Operational templates.
- Explore the W3C XML Schemas of IPS archetype conceptual models with respect to potential practical implementations of the proposed standard.
- Develop a client- server application for testing the openEHR Operational template on an openEHR engine in a local and cloud environment.
- Propose a methodology for transforming a EN ISO 13606 or openEHR archetype conceptual models into a format that enables the creation of archetype instances compatible with NXD.

This paper is divided into sections as follows. In the following section, we review the genesis of the IPS standard and make an overview of the whole standard. A detailed data analysis for the Medication summary section of this standard is presented in Section III. In Section IV we design the IPS with archetype concepts employing well- structured archetype modeling methodology[20]. In Section V, we present our software implementation details. In Section VI, we summarize the obtained results and on this basis a methodology for implementing the Archetype paradigm with the IPS standard is proposed. Section VII summarizes the research results.

II. BACKGROUND

The development of IPS started as a "Patient Summary" (PS) service designed by the epSOS project (2008-2014) [2] to demonstrate the operation of a service infrastructure for cross- border interoperability between systems maintaining electronic health records in Europe. The clinical rationale that led to the definition of the PS dataset is founded on a normative use case scenario of unplanned care (emergency, accident). In this scenario a health professional of country B (country of treatment) gets access to essential information he needs to provide care or consult a patient from country A (country of origin).

By definition [21](p.2) "*An epSOS Patient Summary is a standardized set of basic patient data, which includes the most important clinical facts needed to ensure safe and secure healthcare*". The meaning of the variables included in the proposed PS dataset is described initially in Table 2 of the Guidelines on electronic data exchange (Release 1) [6] adopted by the eHealth Network in 2013. The dataset is split into two groups of data, the *Patient administrative data* group and the *Patient clinical data* group. Both groups distinguish required fields (part of the *Mandatory dataset*) and optional fields (part of the *Extended dataset*) in the PS.

It is noteworthy that from the very beginning the PS document is designed to be compliant to HL7 Common Document Architecture (CDA) Version 2, level 3 [14]. In accordance with this design decision a guide for the implementation of the PS with HL7 technologies has been published [22]. This way the PS document is built by a large number of sections and entry content modules derived from existing HL7 templates and using proprietary encoding. The initial clinical definition of terms in the PS dataset strongly depends on terminology and technical implementations in national infrastructures among European countries. The dataset definition itself lacks a common semantic structure based on standardized data types for information interchange. For example, the proposed translation and transcoding process causes ambiguous identification of a foreign patient. The Guideline on electronic data exchange (Release 2) [7] published in 2016 only partially resolves patient privacy and security issues in the PS dataset design. On the other side this guideline presents a well-structured description of the use case for cross- border PS for unscheduled care and serves as a foundation for the development of the EN 17269 standard of the IPS and its technical specification, FprCEN/TS 17288.

The EN 17269 standard overcomes the major limitations in the implementation of the PS. Its design goal is to be implementation independent; it is all about the IPS data and its logical definition. Unlike the PS, the EN 17269 standard for IPS is underpinned by the System of concepts to support continuity of care (Contsys) EN ISO 13940:2016 [23] and standard data types for information exchange like EN ISO 21090 [15] and HL7 FHIR [8]. For instance, Contsys concepts are used by this standard to interpret the primary use

case of the PS in the epSOS project [5]. Moreover, the technical specification FprCEN/TS 17288 extends the scope of this primary use case to include four scenarios that are combinations of cross- border(international scale) and local (national scale) exchange of PS data with exclusive demands for unscheduled care and scheduled care. References to EN ISO 21090 datatypes support harmonization of the IPS with EN ISO 13606, openEHR, HL7 and FHIR. For example, a Coded element in the IPS standard represents a single concept that provides a reference to a terminology, code system or ontology such as SNOMED-CT [24]. Although the IPS standard employs an abstract data model that is implementation independent such kind of references and Contsys concepts enable semantic interoperability of documents satisfying the requirements of this standard.

An IPS document in cross- border applications must be structured into the following six *Required* sections (Fig. 1) in order to claim full conformance with EN 17269:

- Patient Attributes
- Allergies and Intolerances
- Medication Summary
- Problems
- Provenance
- Cross- Border (conditional)

In national scale scenarios the Cross-border section can be omitted. The three sections with clinical content (Allergies and Intolerances, Medication Summary and Problems) among these six sections are mandatory. In case no information is available for any of these sections, then a dedicated data element in such a section must state the reason for the absence of data. The IPS standard outlines a group of sections that are "required if known", or the so- called group of *Recommended* sections (Immunizations, History of Procedures, Medical Devices and Patient's Address Book). Data in these sections may not be universally available, may not be collected or such data like data in the "Patient's Address Book" section is confidential in some aspect. There is also a group of *Optional* sections (Advance Directives, Functional status, History of Pregnancy, History of Past Illness, Plan of Care, Results, Social History and Vital Signs) that may be omitted in the IPS document without specifying any reason for it. For completeness, the EN 17269 standard allows to include optionally Non-IPS sections in the IPS document as well. The structure and content of such sections is not defined in this standard and it allows to the extend the IPS with condition-specific data.

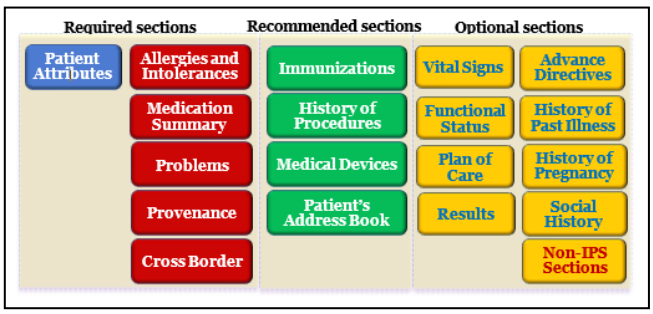


Figure 1. Standard International Patient Summary sections.

Each one of the IPS sections is represented by a hierarchical data structure, where nesting determines the hierarchical level of the data elements. Since the IPS standard is implementation independent the data elements are described in generic form making use of "patterns". Most of these "patterns" correspond to "well- known" data types like Identifier, Coded element, Date- Time and Address. For clarity, the EN ISO 21090 datatypes are referenced to illustrate the expected main characteristics of the data element. Besides, when referring to an EN ISO 21090 datatype it is not assumed that an IPS document implementation must comply with this standard. However, in our opinion compliance of datatypes of the data elements with EN ISO 21090 is a requirement is a necessary condition for ensuring semantic interoperability in cross- border exchange of IPS documents.

Nowadays there exist only two trial implementations of the IPS standard that are reported in the existing literature. Both of them make use of the Message paradigm [17] and implement the IPS standard correspondingly with HL7 FHIR [8] and with HL7 CDA R2 [9]. Both implementation guides share the same design principles and demonstrate the creation of IPS core set of sections as a templated document making use of pre-existing CDA templates and Value sets. An important feature of these sample implementations is the use of terminologies like SNOMED CT [24], ATC [25], LOINC [26] , UCUM [27] for units of measures and EDQM [28] for dose forms and routs of administration.

For example, the HL7 FHIR implementation of a *Medication statement* encodes the *Medication Product Common Name* and *Medication Product code* for the drug SIMVASTATIN in terms of its ATC code C10AA01 (Fig.2).

```
<coding>
        <system value="http://www.whocc.no/atc"/>
        <code   value="C10AA01"/>
        <display value="simvastatin"/>
</coding>
```

Figure 2. Representimg a Coded element in IPS implementations.

The IPS standard does not consider the IPS dataset implementation in the clinical practice. Therefore, these two trial implementations of the IPS serve as a helpful resource for examples about how to interpret the data structure of IPS sections and the "patterns" used to describe the datatypes. The success or the failure of the IPS standard, however, strongly depends on making this standard acceptable for clinical use.

There are several problems that remain to be resolved in order to make the trial implementations useful for the practice. One serious obstacle in this respect is providing user friendly access to terminology services and codes that enable semantic interoperability in cross- border exchange of IPS. Most of these resources are not freely available and identifying the desired code value is time consuming. Some other resources like the set of EN ISO IDMP set of standards [29] are not fully implemented in practice. Although the current IPS standard does not impose obligatory usage of terminology services and

codes, it is self-understood that semantic interoperability is impossible without binding clinical concepts and terms to a coding system. Another major problem that concerns the clinical practice is the management of IPS records. The trial implementations illustrate merely the interpretations of IPS data making use of HL7 related technologies that represent the view of the Message paradigm in Health informatics.

At first glance it seems to be straightforward to accomplish mapping of datatypes between Message paradigm technologies and Archetype paradigm technologies because more or less these datatypes derive from EN ISO 21090. However, each one of these two major groups of technologies has their inherent limitations in terms of management of EHR like IPS data [30] [31] [32] [33]. This refers to storage and exchange of IPS data taking in consideration related issues of privacy and confidentiality regulations, language translation of clinical concepts and differences in national legislation as well as availability of tools for executing queries and data visualization. Finally, there should be a convenient way for transforming existing heterogeneous clinical data into a document structure that conforms to the IPS standard. In short, the available IPS implementation guides are too simplified and biased by the advantages and disadvantages of one of the existing paradigms in health informatics for exchange of clinical information.

In the following sections we consider the IPS design from the perspective of the Archetype paradigm. For this purpose, we first analyze the data structure of a required section in the IPS standard. Next, we will implement this section in terms of archetype concepts. Unlike the existing IPS implementation samples we won't just create a computer model of an IPS section. From practical point of view, it is important to investigate and discuss management of instances of the thus created IPS model. Therefore, we build a server- client web application and demonstrate IPS data management of archetype with this application.

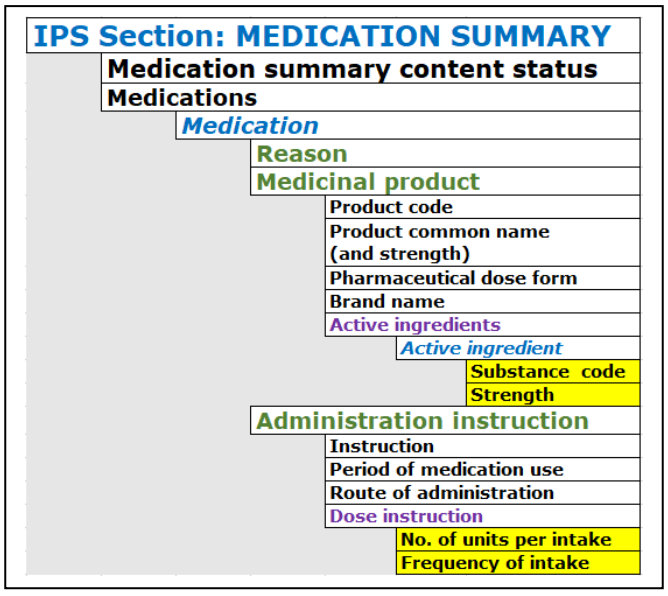


Figure 3. The Medication Sumamry required structure in EN 17269.

III. DESCRIPTION OF THE DATA MODEL

Without loss of generality, we consider the *Medication Summary* section of the IPS standard, which is one of its six *Required* sections (Fig. 3).

The hierarchical data structure and the data types used in this section are described in full details in Tables 15 and 16 of the EN 17269 standard (Fig. 3). The final draft version of this standard describes the same way the rest of the sections of the IPS, where the dataset borrows data types from EN ISO 21090. Therefore, the same approach for can applied for analyzing the remaining sections of the IPS standard. The standard provides detailed description with all the information necessary for building a conceptual model of these sections in terms of archetype concepts.

The "IPS Section: MEDICATION SUMMARY" is closely related to EN ISO IDMP set of standards and it is important part of European eHealth projects like the epSOS project. The IPS standard taking into consideration the limited scope of EN ISO IDMP implementations has relaxed the requirement for using IDMP notations and specifications by introducing high level of abstraction in concept descriptions. This way the IPS standard leaves the option for adopting IDMP notation and specification with *Medication summary* concepts, when they become available.

The dataset of the *Medication summary* section comprises a *Content status* data element and a list of *Medications*. For clarity we keep the same names and order of introducing the data elements of the *Medication Summary* dataset as they are defined in the IPS standard. The *Content status* element is of type denoted in the IPS standard using the pattern *Coded Element*. The semantic meaning of this pattern with reference to *Content status* is a predicate taking for example, values true and false. When the *Content status* is *true*, then the list of *Medications* shall be not empty. Otherwise, the list of *Medications* shall be *empty*.

Each one of the elements of the *Medications* list contains:

- *Reason*
- *Medicinal product*
- *Administration instruction*

The *Reason* element is *Optional* in the *Medications* list. It has the purpose to state, if known or allowed by the patient, the reason for the *Medication* prescription.

The *Medicinal product* and the *Administration instruction* are *Required* datasets of elements that together provide a "*minimal and non- exhaustive*" description of that product. The following elements belong to the *Medicinal product* dataset:

- *Product code*
- *Product common name(and strength)*
- *Pharmaceutical dose form*
- *Brand name*
- *Active ingredients*

The *Product code* element is *Optional*, and its datatype is described with the pattern *Coded element*. The semantics of the product code description in the IPS standard implies a reference to some unique identifier from the EN ISO IDMP set of standards. However, the IPS does not impose such a requirement rather allows for using a more general code from

a terminology like ATC as shown in the sample implementation on Fig. 2.

The *Product common name* (*and strength*) is a *ReqiuredIfKnown* element of type *String*. According to the IPS standard it is used to record in free text a non-proprietary name of the pharmaceutical product eventually including the strength of each ingredient. It should be consistent and associated with the *Product code* value. For example, once the *Product code* is selected from a given terminology classification (ATC) then this determines the *Product common name* ("*simvastatin*") (Fig. 2).

The *Pharmaceutical dose form* is a *Required* element described by a value described by the pattern *Coded element.* This element denotes the physical state of Medication (*Solid*, *Liquid*, *Semi- Liquid* and their variants). The codes for these states are listed in Table A.1 of EN ISO 11239, which is one of the EN ISO IDMP group of standards.

Brand name is an *Optional* element of type *String*. For a given *Product code* there could be multiple titles under which the medication is sold on the market. For example, given the ATC code C10AA01 the *Brand name* could be "*Simvastatin 40 mg Filmtabletten*" or "*Simvastatin 40 mg Tablets*". The standard recommends providing a *Brand name* always when it is justified by the healthcare professional or the medicinal product is of biological origin.

The list of *Active ingredients* is represented by a set of the substances that separately or in combination produce the intended effect of applying the medicinal product. This list is a *Required* element in the Medicinal product dataset. The elements of the list are described in terms of *Substance code* and *Strength*.

Substance code is a *Required* element described by a value described by the pattern *Coded element.* The IPS standard refers to codes in conformance with the EN ISO IDMP set standards and in particular, satisfying the information model for substance codes EN ISO 11238:2018 like the Chemical Abstract Service (CAS) Registry Numbers, European Inventory of Existing Commercial Chemical Substances (EINECS), European Drug Codes (XEVMPD) and Japanese Drug Codes. The sample implementation of IPS by HL7 FHIR stores the *Substance code* in terms of its RxNorm [34] code and respective name (Fig. 4).

```
<coding>
    <system >
        http://www.nlm.nih.gov/research/umls/rxnorm
    </system/>
    <code    value="36567"/>
    <display value="Simvastatin"/>
</coding>
```

Figure 4. Sample representation of Substance code in the IPS standard.

Currently, the IPS standard allows using "*well- known*" names for ingredients instead of IDMP or other terminology codes, for example, "*paracetamol*".

The *Strength* of the *Active ingredient* is a *Required* element described by a value described by a pattern termed as *Ratio.* When the IPS standard explains this pattern, it makes a reference to the RTO type in the EN ISO 21090 standard having numerator and denominator of type QTY (quantity), where the denominator shall be nonnegative (Fig. 5).

```
<strength>
    <numerator>
        <value  value="40"/>
        <unit   value="mcg"/>
        <system value="http://unitsofmeasure.org"/>
        <code   value="mg"/>
    </numerator>
    <denominator>
        <value  value="1"/>
        <unit   value="tablet"/>
        <system value="http://unitsofmeasure.org"/>
        <code   value="{tablet}"/>
    </denominator>
</strength>
```

Figure 5. Sample representation the Strength of the Active ingredient.

The *Administration instruction* dataset comprises the following data items:

- *Instruction*
- *Period of medication use*
- *Route of administration*
- *Dose instruction*

Instruction is an *Optional* element of type *Text* providing textual information about the procedure to applying the *Medication*. By default, it is used to summarize the content in the remaining three data elements of *Administration instruction.*

Period of medication use is an *Optional* element of type described by the pattern *Period* in the IPS standard. The period is a general concept that could be implemented in terms of start and end dates or in terms of "floating" time for duration ("*5 days*", or "*1 month*").

The *Route of administration* is an *Optional* element described by the pattern *Coded element.* For some reason a description for this element is missing for the *Medication summary* in the IPS standard. However, a clear description for it is provided in the EN ISO 11239 standard, namely, "*path by which the pharmaceutical product is taken or makes contact with the body*". Moreover, Table A.9 in the same standard provides a list with examples for some of the most frequently used *Route of administration* values.

Finally, *Dose Instruction* is a *Required* component of the *Administration instruction* dataset comprising two data elements, *Number of Units per intake* and *Frequency of intake*. Both data elements are *Required* and described correspondingly by patterns *Range or Quantity* and *General Time Specification*. The *Number of Units per intake* is represented by a numeric value and a measurement representation for the unit ("40 mg", "2 tablets"). In a similar way the *Frequency of intake identifies* the number of intakes per a specified period (hour, day, week, month).

This detailed understanding of the data model is an important prerequisite for representing it with archetype concepts in the following section of this paper.

IV. IPS Design with Archetype Concepts

In this section we explore the design of a model of the IPS *Medication Summary* in terms of the Archetype paradigm. Additionally, compare the models of the IPS *Medication Summary* obtained by employing respectively, EN ISO 13606 archetypes and openEHR archetypes in the design process. To achieve this objective, we follow a five- stage methodology for archetype design [20]. It allows us to transform correctly a clinical document like the IPS into an archetype conceptual model after the detailed analysis of its data model in the previous section. Some of the most popular software tools for creating an archetype model are LinkEHR Studio [35], the Template Designer [36] and the Archetype Editor [37]. The Template Designer and Archetype Editor are more specialized and appropriate to employ with the openEHR AOM, while the LinkEHR Studio is more suitable for creating a EN ISO 13606 AOM.

These tools have the potential to bind semantic context to the archetype conceptual model from terminology servers and this way map terms to international standards like LOINC, ATC, SNOMED-CT, RxNorm, and ICD [38]. It allows the IPS conceptual model to deliver meaningful, reliable, semantical clinical information at the point of care. At this stage, however, the IPS standard does not impose requirements to use multilingual international reference terminologies and in our research, we have followed strictly the text of this standard. For clarity, we preserve the names of the concepts, the semantics of the data types employed correspondingly with EN ISO 13606 and openEHR as well as the constraints on occurrences of the data elements the way they are specified in the final draft of the IPS standard.

The design of a clinical document with archetype concepts requires good understanding of the dual- level information model in ISO 13606 and openEHR. This information model comprises a Reference model and Archetype model (Fig. 6).

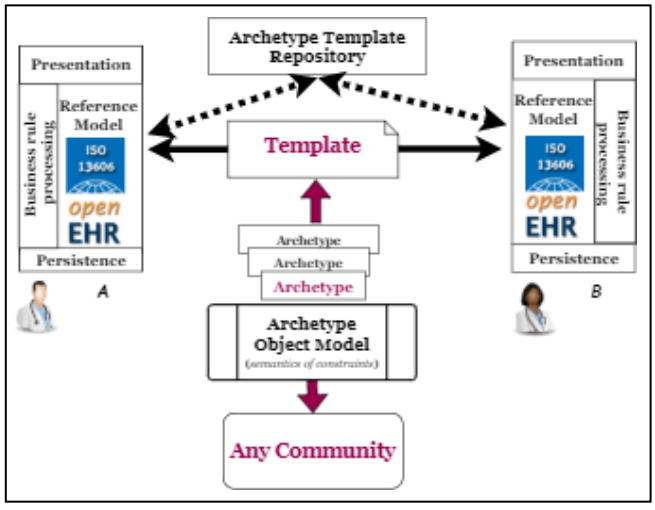


Figure 6. W3C XML Schema of an IPS Section in EN ISO 13606 AOM.

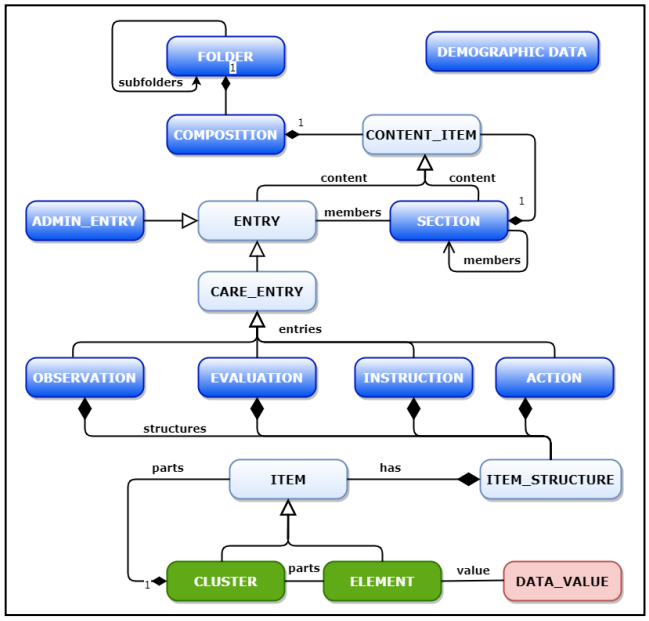


Figure 7. UML class diagram of the openEHR Reference Model.

Without loss of generality we shortly describe the Reference Model of openEHR because in the following section we will present a web application making use of this model. It is displayed in Fig. 7, where for clarity the diagram is simplified by omitting some classes like LOCATABLE and PATHABLE, corresponding to RECORD_COMPONENT in EN ISO 13606. The classes displayed with dark background are concrete classes and the rest are abstract classes. Once a patient creates an EHR or IPS document in particular, this patient is associated with a unique UID identifier. All references to openEHR patient data are made by means of this UID. Committed changes of the document are considered as a contribution and such contributions are represented by a new instance of class COMPOSITION. All contributions are identified by a version number related to the UID of the EHR. This way a FOLDER instance serves as a container for a versioned structure of COMPOSITION instances (contributions for a given EHR). The COMPOSITION instances contain all the information in the clinical document which is the IPS in our case. The base definition of class COMPOSITION needs to be enriched in each given case of EHR with semantic constraints specific to this case. This happens by creating an archetype of type COMPOSITION in the Archetype Object Model. Further on, archetypes of COMPOSITION type are organized in Templates that facilitate reusability of archetypes. It is noteworthy that the latest edition EN ISO 13606-2:2019 has introduced templates as part of its archetype object model as well. A Template allows to dynamically update the semantic structure of the clinical document by adding slots for splitting the COMPOSITION archetype into multiple SECTIONs, which can be further on structured into archetypes, representing semantic constraints determined by the clinical practice as OBSERVATION, EVALUATION, INSTRUCTION and ACTION. The data ITEMs are in the leaves in this tree

structure, where concrete data ELEMENTs may exist standalone or organized in CLUSTER groups (list or tree). The data types assigned to each ELEMENT derive from class the abstract class DATA_VALUE. The openEHR datatypes derive from EN ISO 13606.

All of these allow us to make the conclusion that the IPS can be designed according to the dual- level information model in EN ISO 13606 and openEHR. This model allows not only to reproduce the structural constraints in the IPS document. It enables semantic interoperability, flexibility and reuse of the archetype concepts. Instead of using predefined XML schemas (templates) in HL7 CDA and HL7 FHIR the dual- level information model allows to develop custom archetype concepts that accommodate the specific needs in the clinical practice. A great advantage in using the Archetype paradigm is the introduction of a common information model for validation the template instances in EN ISO 13606 and openEHR.

The IPS standard represents this clinical document as a tree hierarchical of sections, where each section is composed of datasets that can be designed as clusters (lists or trees) of data elements. The IPS *Medication Summary* data model is a typical example of this kind. A Mind map of this data model designed with archetype concepts from EN ISO 13606 allows to visualize the structure of the conceptual model as a single archetype (Fig. 8). This model maps exactly the hierarchical structural constraints of the IPS Medication summary as described in the IPS standard to semantically correct structures within the archetype model.

At the highest level we observe the CONTENT_STATUS (*Content status)* and the LIST_OF_MEDICATION (*Medications)*, where each element MEDICATION of this list is a tree structure of REASON (*Reason)*, MEDICINAL_PRODUCT (*Medicinal product)* and ADMINISTRATION_INSTRUCTION (*Administration instruction).* For comparison purposes, the names of the respective data elements in the IPS are placed in brackets. All the occurrences, cardinalities and data types of the elements in the leaves of this hierarchical structure are assigned in accordance with the semantic interpretation of the patterns in the IPS standard. These settings are denoted on Fig. 9, where the archetype models correspondingly, in EN ISO 13606 (left side) and openEHR (right side), are displayed.

In technical terms, the creation of an archetype model with openEHR first requires to design with the Archetype Editor an archetype a type COMPOSITION archetype (representing the whole IPS) with slots for SECTION archetypes for each of the sections of the IPS, create the SECTION archetypes themselves (representing each of the sections of the IPS) as well as specialized ENTRY archetypes (representing the CONTENT_ITEM in each one of the IPS sections) like OBSERVATION, EVALUATION, INSTRUCTION or ACTION. For example, the contents of the *Medication Summary* section have been created as an INSTRUCTION specialization of the ENTRY archetype because medication orders or multi- drug courses in the clinical practice instructions are considered as instructions. Next, a template can be created with the Template Designer, where all the above archetypes get assembled into a single archetype model of the IPS. Once the constraints of the Reference model are satisfied this template to be exported as Operational template to a openEHR platform for managing instances of the thus created archetype model. The procedure is similar in the case of creating a EN ISO 13606 archetype model with LinkEHR Studio.

The conceptual design of the IPS section of the archetype models can be explored in W3C XML Schema format as well. This approach to investigate the IPS conceptual model is a novelty in the existing literature because a W3C XML Schema Definition Language (XSD) model of the IPS is not presented in the proposed prEN 17269 standard. At the same time, its practical implementation implies the use of web services, where the XSD model specification is important. This model is suitable for implementations in NXD, where XQuery can serve as a good replacement for the absence of an Archetype Query Language in the AOM of EN ISO 13606. For comparison, the XSD model of openEHR concepts includes a lot of metadata payload. Real- life IPS applications process large numbers of openEHR XSD model instances.

Finally, we note that in the existing literature there is evidence that openEHR archetypes can be transformed to EN ISO 13606 archetypes making use of the common ISO 21090 data set [39] The mapping from EN ISO 13606 to openEHR is not explored so far. As a side result, we have established that it is not possible to convert EN ISO 13606 archetypes directly into openEHR archetypes.

The reason is that the concrete class ENTRY from the Reference model of EN ISO 13606 is mapped to the abstract class ENTRY in the Reference model of openEHR. Therefore, we can export a EN ISO 13606 archetype into a valid openEHR Operational template. However, it is not possible to create instances of that template. In openEHR there are concrete specializations of class ENTRY like OBSERVATION, EVALUATION, INSTRUCTION and ACTION. Hence, it is very difficult to map a concrete ENTRY class in EN ISO 13606 to some of these specializations of the openEHR class ENTRY.

Once the openEHR archetype model is built the conceptual model can be exported to valid openEHR Operational template. Thus, in contrast to the existing implementation guides of the IPS, which merely publish XML schemas adapted to the IPS standard, we demonstrate the obtained archetype model in action. In the following section we show results of execution and testing the obtained openEHR conceptual model in a client- server application making use of an openEHR engine for processing instances of the Operational template for the *Medication Summary* section. Without loss of generality the whole IPS can be designed with archetype concepts in the above described way.

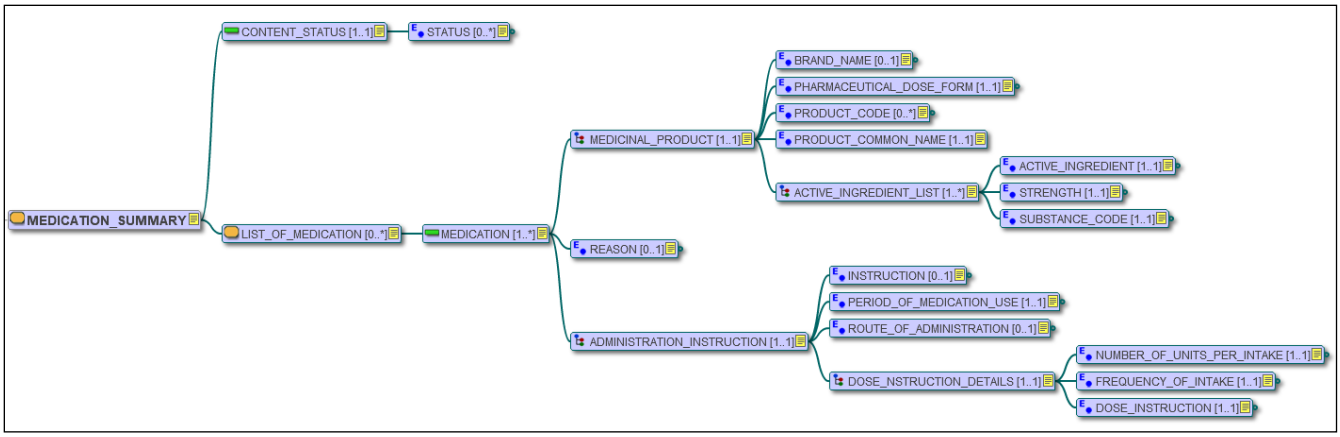


Figure 8. Mind map design of the IPS Medication Summary Section in EN 17269:2018 with a EN ISO 13606 archetype.

V. SOFTWARE IMPLEMENTATION

The software implementation of the IPS conceptual models in a client server application has been developed with the objective to evaluate the applicability of the Archetype paradigm in implementing the proposed IPS standard. For this purpose, the openEHR Operational template of the *Medication Summary* section has been installed and run on both a local CaboLabs openEHR platform [36] and a cloud-based Code4Health platform [40]. Here we present results from a local installation of an openEHR server.

Initially, the Operational template for the *Medication Summary* section is uploaded on the openEHR server and an EHR UID is created for a given patient. On Fig. 10 there are two patients registered with their UIDs. The different versions of updates of the IPS for the second patient in the list are displayed as his Contributions on the same figure.

These Contributions are created and managed by a client web application. This web application employs the RESTful API published by the openEHR server to manage instances of the Operational template that actually are XML documents validated with respect to the information model of openEHR.

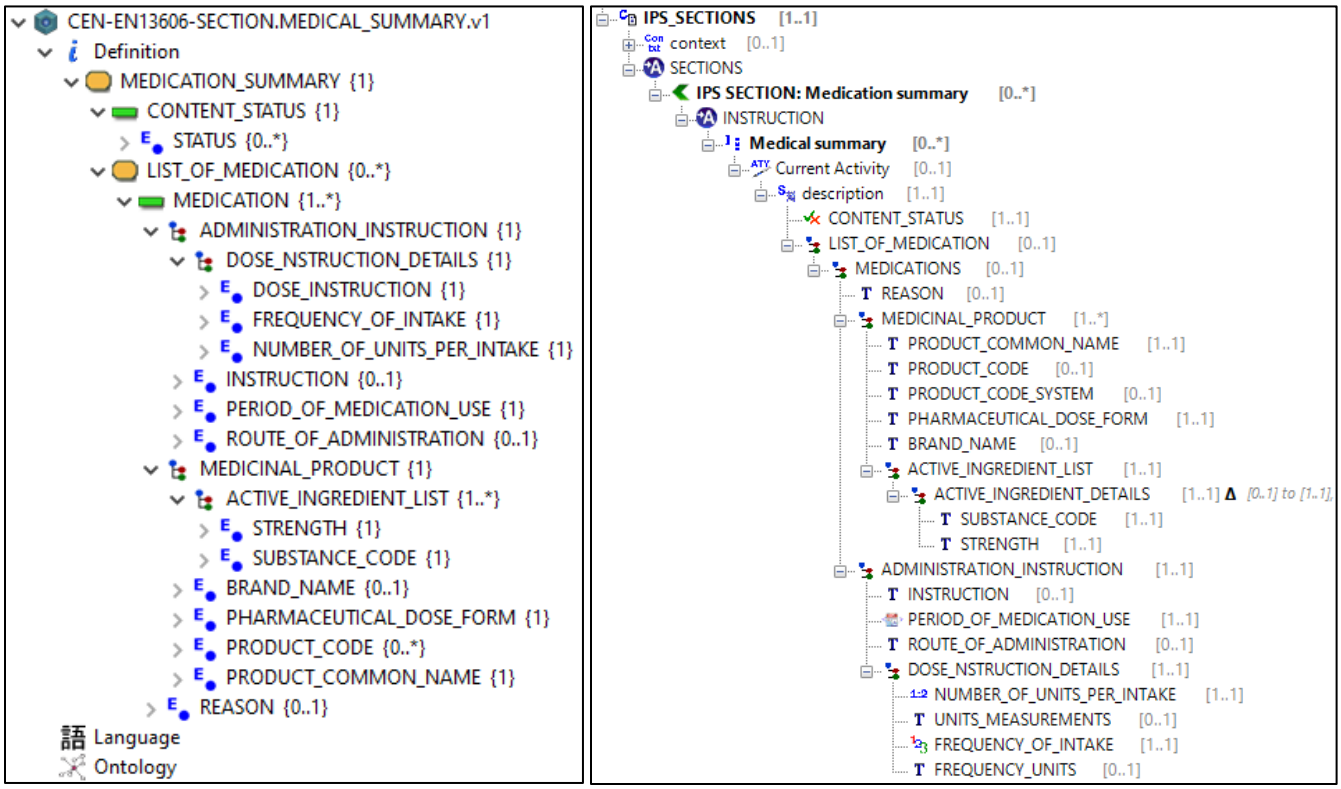


Figure 9. IPS Medication Summary design as a EN ISO 13606 archetype (left side) and as a template of openEHR archetypes (right side).

#	UID	Date
1	bb2dea41-e1dc-415b-95ea-7693d5c14c20	2020-02-29 12:09:31
2	d76df9d8-3013-43ac-9dbd-0c446c4a67d8	2020-02-29 16:09:50

Contributions for EHR

#	UID	Date
1	e87044b6-d2ce-4528-8473-db3e8aa5401c	2020-02-29 16:10:12
2	14a0ba7b-e591-4158-b2d1-38e53e10cc5b	2020-02-29 19:17:03
3	6a76d4be-e034-4080-b8f7-12cd04587aee	2020-02-29 19:41:42

Figure 10. Registered EHR UIDs and versioned Contributions.

The client application at this stage operates in two modes, "writer" (update, add or delete content) and "reader" (read only content). Snapshots of the client application are shown in Fig. 12 and Fig.13.

The graphical user interface of the client is custom designed in PHP. The combo box on the right upper side allows to load an existing contribution or create blank form by the patient, whose EHR UID is displayed below the title of the form. The combo box on the left upper side allows to select an existing Medication in the Medication summary section and display details for it in the fields of the form. Currently, for the registered patient we demonstrate in Fig. 12 and Fig.13 the contents of two Medication. One of them is the same drug, SIMVASTATIN, used in the sample implementation guide for the IPS produced by HL7 FHIR. With this we want to demonstrate that the conceptual model implemented with archetype concepts can handle the same use cases of the IPS as those included in the existing in the literature. For example, the input on Fig. 13 for *Product code* (PRODUCT_CODE on Fig. 9, right side) and *System* the (PRODUCT_CODE on Fig. 9, right side) match completely those input in the sample HL7 FHIR template. Besides, the data elements holding these input values are part of the archetype model and respectively, they are effectively exchanged and managed in the communication between the client and the openEHR server. Note that the input for PRODUCT_CODE is required in the IPS standard. However, this way we show that the obtained archetype model can be extended to serve bindings with terminology specifications and notations. For clarity, we have extended the model by adding input identifying the measures for units and frequency when recording *Dose instructions*. Moreover, we have used data published in EN ISO 11239 to populate with standard coded terms the combo boxes for *Pharmaceutical dose form* and *Route of administration*. Additionally, the Medication summary form provides default values, where possible and a calendar wizard for input of dates. The objective in the client form design has been to enhance the user experience in dealing with IPS.

The buttons in the bottom of the form allow adding a new *Medication* to the list of *Medications* in the IPS as well as new *Active ingredient* to a *Medication*. The web application is fully functional and demonstrates that contents of the IPS can be exchanged employing an information model designed with archetype concepts. Further on IPS exchange can be extended to take place between openEHR servers serving as access points and this way enable true cross- border exchange of IPS data. As far as we know this sample web application is the first operational sample implementation of the IPS standard.

VI. DISCUSSION

In this paper, we have evaluated the applicability of the Archetype paradigm in software applications that implement the draft version of the IPS standard. Although the Archetype paradigm is not in the focus of this standard, we have demonstrated that it can be successfully employed in software applications.

The practical experience in producing the here reported results allow us to propose a methodology for applying the proposed IPS standard in use cases where semantic interoperability is a requirement (Fig. 11).

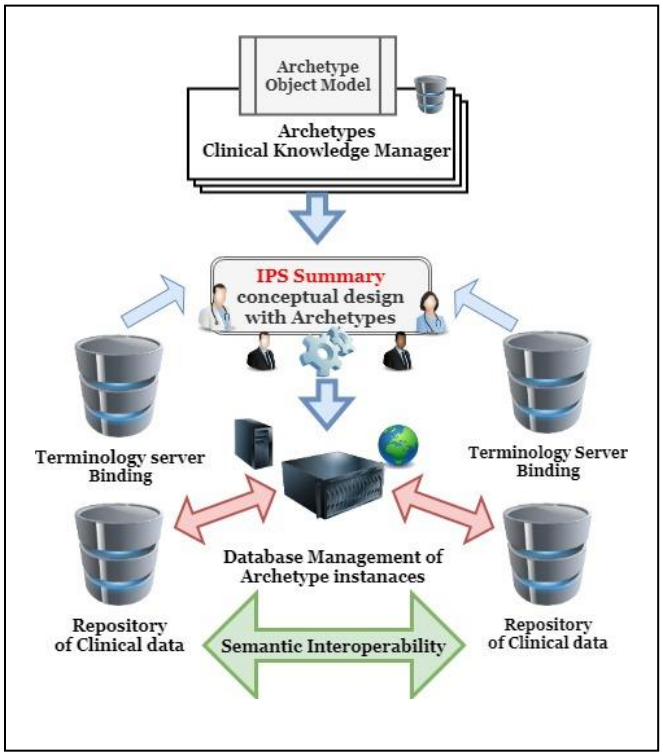


Figure 11. Methodology for semantic interoperability of the IPS.

The first stage in this methodology is to create a correct conceptual model of the IPS document. The archetypes in this model serve as "*plug-and-play*" building blocks for semantic interoperability that can be imported from a repository known a Clinical Knowledge Manager. At this stage we just note to the important requirement for binding of the archetype model to terminology servers for specifications and notations. Therefore, it would be useful by the next version of the IPS standard to obtain well organized and structured public access not only to the promised standard set of SNOMED-CT code set, but also access to other registers for international codes, specifications and terms mentioned in this standard. The obtained archetype conceptual model is ready for reuse in software applications. Clients select an archetype conceptual

model and manage instances of that model in terms of semantic interoperability. Unlike the Messaging Paradigm the implementation of the Archetype paradigm with the proposed IPS standard makes possible all data, information and knowledge in each system to be available in a uniform and standard way.

VII. Conclusion and Future Work

This paper investigates the design and software implementation of the International Patient Summary according to the final draft versions of EN 17269 and FprCEN/TS 17288. These versions make use exclusively of HL7 FHIR and HL7 CDA technologies with the purpose to develop a single, common specification of the minimal and non- exhaustive set of clinical data that can be used by all clinicians for cross- border, unscheduled care of a patient.

In this paper we have presented a detailed analysis of the dataset of the IPS standard with a focus on the IPS Medication summary section. The design objectives of this standard are to make it implementation- independent. It attempts to harmonize requirements among the existing standards and specifications. However, our analysis of the genesis of IPS standard and existing sample implementation guides shows that it is biased by the Message paradigm information model and data set. Therefore, in this paper, we explore how the objectives of this standard can be achieved by applying the Archetype paradigm approach. Especially because the Archetype paradigm is implemented by the EN ISO 13606 standard which is the leading standard for health information exchange approved in the European Union. In the literature it is recognized that reusability and semantic interoperability of clinical data like the IPS sections are delivered essentially in terms of Archetype concepts. Since the IPS sections in EN 17269 have quite a similar structure then the here obtained results can be extended and design the whole IPS in terms of EN ISO 13606 and openEHR archetype concepts. Both archetype models support binding to semantic context provided by terminology servers like SNOMED-CT.

We have established that the EN ISO 13606 AOM cannot be used to create instances compatible with the openEHR AOM. It is an important conclusion because the compatibility of transformation from EN ISO 13606 to openEHR has not been confirmed in the existing literature. The numerical experiments demonstrate that both archetype models can be exported in W3C Schema definitions, where the XSD of openEHR AOM contains considerably larger payload of metadata. Therefore, it is preferable to manage instances of openEHR archetype models by means of operational templates on native openEHR platforms or relational databases with API support for publishing web services.

It is noteworthy that the current versions of the IPS standard and its technical specification do not consider an XSD model of the IPS. At the same time, practical implementations of this standard rely on web services where the specification of the XSD models is important. We have demonstrated this approach in a client server application with real medicinal data and terminology codes, where the same operational template of the IPS section can run both on a local and on a cloud-based openEHR engine. Implementation details are reported in related research work [41] [42] Accordingly, instances of EN ISO 13606 archetypes can be managed in NXD as it is demonstrated in the use cases of FprCEN/TS 17288 with HL7 concepts.

In our future work, we first plan to explore the feasibility of executing an EN ISO 13606 archetype model of the IPS on a NXD. At the same time, we will try to improve the user interface experience for managing the IPS in client applications. In addition, we plan to create a complete Archetype model of all the sections in the IPS standard. Finally, we will continue work directed on exchanging bulk data between servers that provide support for EN ISO 13606 or openEHR.

In summary, the obtained results are presented in a uniform methodology for implementing the IPS in terms of the Archetype paradigm. These results are novel because the Archetype paradigm is not considered in the final draft version of the IPS standard and implementation guides in the existing literature. They serve to extend the practical experience in cross-border sharing of clinical data represented in terms of semantic interoperability of archetype concepts.

Acknowledgment

This research is supported by the National Scientific Program eHealth in Bulgaria.

References

[1] E. Krastev, D. Tcharaktchiev, L. Kirov, P. Kovatchev, S. Abanos and A. Lambova, "Software Implementation of the EU Patient Summary with Archetype Concepts," in GLOBAL HEALTH 2019, The Eighth International Conference on Global Health Challenges, Porto, Portugal, 2019.

[2] European Patients Smart Open Services (epSOS), "D3.2.2 Final definition of functional service requirements- Patient Summary," 29 October 2012. [Online]. Available: https://ec.europa.eu/digital-single-market/en/news/cross-border-health-project-epsos-what-has-it-achieved. [Accessed 5 June 2019].

[3] A. G. Ekeland and G. Ellingsen, "Assessing an Electronic Health Record (EHR): How Do Basic Assumptions in Traditional Health Technology," International Journal on Advances in Life Sciences, vol. 9, no. 1 and 2, pp. 1 - 10, 2017.

[4] European Commission, "Joint Action to support the eHealth Network [JAseHN][677102] - Joint Actions," 30 June 2018. [Online]. Available: https://webgate.ec.europa.eu/chafea_pdb/health/projects/677102/summary. [Accessed 5 June 2019].

[5] eHealth DSI, "PS Use Case," 30 April 2019. [Online]. Available: https://ec.europa.eu/cefdigital/wiki/display/EHOPERATIONS/PS+Use+Case. [Accessed 4 June 2019].

[6] eHealth Network, "Guidelines on Minimum/Nonexhaustive Patient Summary dataset for Electronic Exchange in Accordance with Cross- Border Directive 2011/24/EU. Release 1.," European Commission, 19 November 2013. [Online]. Available: http://www.ehgi.eu/Download/eHealth Network%204 [Brussels 19-Nov-2013] Guidelines_Patient_Summary_Cross_Border_en.pdf. [Accessed 22 February 2020].

[7] eHealth Network, "Guideline on the electronic exchange of health data under Cross-Border Directive 2011/24/EU. Release 2.," 21 November 2016. [Online]. Available: https://ec.europa.eu/health/sites/health/files/ehealth/docs/ev_20161121_co10_en.pdf. [Accessed 22 February 2020].

[8] HL7 FHIR, "FHIR Release 4. International Patient Summary Implementation Guide," 27 December 2019. [Online]. Available: https://build.fhir.org/ig/HL7/fhir-ips/. [Accessed 5 June 2019].

[9] HL7 International, "HL7 CDA R2. International Patient Summary Implementation Guide," 2018. [Online]. Available: http://international-patient-summary.net/mediawiki/index.php?title=IPS_implementation guide_1. [Accessed 4 June 2019].

[10] CEN/TC 251, "EN 17269. Health informatics - The Patient Summary for Unscheduled, Cross- border Care," European Committee for Standardization, Brussels, 2019.

[11] CEN/TC 251, "FprCEN/TS 17288. Health informatics - The International Patient Summary: Guidance for European Implementation," European Committee for Standardization, Brussels, 2018.

[12] openEHR Foundation, "10. Archetypes and Templates," openEHR, December 2018. [Online]. Available: https://specifications.openehr.org/releases/BASE/latest/architecture_overview.html#_archetypes_and_templates. [Accessed 4 May 2019].

[13] CEN/TC 215, "ISO 13606-1:2008 Health informatics -- Electronic health record communication -- Part 1: Reference model," 2 2008. [Online]. Available: https://www.iso.org/standard/40784.html. [Accessed 6 April 2019].

[14] HL7 International, "HL7 Version 3 Clinical Document Architecture," 2019. [Online]. Available: http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7. [Accessed 5 June 2019].

[15] ISO/TC 215 Health informatics, "ISO 21090:2011 Health informatics -- Harmonized data types for information interchange," Feb 2011. [Online]. Available: https://www.iso.org/standard/35646.html. [Accessed 8 Apr 2019].

[16] ART-DECOR, "ART-DECOR Open Tools," 2019. [Online]. Available: http://art-decor.org/art-decor/decor-project--hl7ips.

[17] G. Freriks, G. de Moor and D. Karla, "White paper: Archetype paradigm: an ICT revolution is needed," 13 March 2007. [Online]. Available: http://www.eurorec.org/files/filesPublic/ArchetypeParadigmFeb2007.pdf. [Accessed 5 June 2019].

[18] ISO/TC 215, "ISO 13606-1:2019 Health informatics -- Electronic health record communication -- Part 1: Reference model," ISO, 2019. [Online]. Available: https://www.iso.org/standard/67868.html. [Accessed 6 June 2019].

[19] ISO/TC 215, "ISO 13606-2:2019. Health informatics -- Electronic health record communication -- Part 2: Archetype interchange specification," ISO, 2019. [Online]. Available: https://www.iso.org/standard/62305.html. [Accessed 5 June 2019].

[20] D. Moner, J. A. Maldonado and M. Robles, "Archetype modeling methodology," Journal of Biomedical Informatics, vol. 79, pp. 71-81, 2018.

[21] epSOS project, "The epSOS Patient Summary," 2014. [Online]. Available: http://ec.europa.eu/newsroom/document.cfm?action=display&doc_id=723. [Accessed 20 February 2020].

[22] G. Cangioli, C. Gessner and K. Hyppönen, "EXPAND. epSOS Patient Summary, ePrescription, eDispensation and Common Modules HL7 CDA R2 Implementation Guide," HL7 Foundation, GEMATIC, THL, 2015 Sept 2015. [Online]. Available: https://ec.europa.eu/cefdigital/wiki/download/attachments/35201366/WP3A_epSOS_EED_PSePeD_CM_CDAIG_1_1.pdf. [Accessed 20 Feb 2020].

[23] CEN/TC 251, EN ISO 13940:2015. Health informatics-System of concepts to support continuity of care, Comite Europeen de Normalisation (CEN), 2016.

[24] SNOMED International, "SNOMED CT," 2019. [Online]. Available: http://www.snomed.org/snomed-ct/get-snomed-ct.

[25] WHO Collaborating Centre for Drug Statistics Methodology, "ATC classification system. ATC/DDD Index 2020," 2020. [Online]. Available: https://www.whocc.no/atc_ddd_index/. [Accessed 22 February 2020].

[26] Regenstrief Institute, Inc, "LOINC. The international standard for identifying health measurements, observations, and documents," 2020. [Online]. Available: https://loinc.org/. [Accessed 24 February 2020].

[27] US National Library of Medicine, "Unified Code for Units of Measure (UCUM)," 2020. [Online]. Available: https://ucum.nlm.nih.gov/. [Accessed 22 February 2020].

[28] C. o. Europe, "EDQM. Standard Terms database," 2020. [Online]. Available: https://standardterms.edqm.eu/.

[29] European Union agencies network, "Data on medicines (ISO IDMP standards): Overview," European Medicines Agency, 2020. [Online]. Available: https://www.ema.europa.eu/en/human-regulatory/overview/data-medicines-iso-idmp-standards-overview. [Accessed 22 February 2020].

[30] J. L. C. De Moraes, W. L. De Souza, L. F. Pires, L. T. Cavalini and A. F. Do Prado, "Using the dual-level modeling approach to develop applications for pervasive healthcare.," Journal of Mobile Multimedia, vol. 9, no. 1–2, p. 111–127, 2013.

[31] T. Beale, "Standards Classification," 8 September 2008. [Online]. Available: https://openehr.atlassian.net/wiki/spaces/stds/pages/5373958/Standards+Classification. [Accessed 2 June 2019].

[32] P. Schloeffel, T. Beale, G. Hayworth, S. Heard and H. Leslie, "The Relationship between CEN 13606, HL7, and OpenEHR," in HIC 2006 and HINZ 2006:Proceedings, Health Informatics Society of Australia, 2006.

[33] B. Blobel, "Interoperable EHR Systems – Challenges, Standards and Solutions," European Journal for Biomedical Informatics, vol. 14, no. 2, pp. 10-19, 2018.

[34] N. L. o. Medicine, "RxNorm," National Library of Medicine, 2020. [Online]. Available: https://www.nlm.nih.gov/research/umls/rxnorm/index.html. [Accessed 11 February 2020].

[35] VeraTech for Health, "LinkEHR Interoperability Platform," 2019.

[36] P. P. Gutiérrez, "Towards the Implementation of an openEHR-based Open Source EHR Platform (a vision paper)," in MEDINFO 2015: EHealth-enabled Health: Proceedings of the 15th World Congress on Health and Biomedical Informatics, São Paulo, Brazil, 2015.

[37] Ocean Health Systems, "Archetype Editor," 2015.

[38] C. Coimbra, M. Esteves, F. Miranda, F. Portela, M. Santos, J. Machado and A. Abelha, "Improving the Codification of Hospital Discharges with an ICD-9-CM Single-page Application and its Transition to ICD-10-CM/PCS," International Journal on Advances in Life Sciences, vol. 10, no. 1&2, pp. 23 - 30, 2018.

[39] T. Beale, "openEHR to ISO 13606-1, ISO 21090 mapping," 15 May 2015. [Online]. Available: https://openehr.atlassian.net/wiki/spaces/stds/pages/5373954/openEHR+to+ISO+13606-1+ISO+21090+mapping. [Accessed 5 May 2019].

[40] Code4Health, "Code4Health Platform," 2019. [Online]. Available: Code4Health Platform. [Accessed 10 Apr 2019].

[41] D. Tcharaktchiev, E. Krastev, P. Petrossians, S. Abanos, H. Kyurchiev and P. Kovatchev, "Cross-border Exchange of Clinical Data using Archetype Concepts Compatible with the

International Patient Summary," in Digital Personalized Health and Medicine. Proceedings of MIE 2020, vol. 270, L. Pape-Haugaard, C. Lovis, I. Cort Madsen, P. Weber, P. H. Nielsen and P. Scott, Eds., Geneva, Studies in Health Technology and Informatics, IOS Press, 2020 (in print).

[42] P. Kovachev, "openEHR implementation of International Patient Summary based on the EN 17269 standard," eHealth National Scientific Program of Bulgaria, 2020. [Online]. Available: https://www.youtube.com/watch?v=lDReBt9DsAs. [Accessed 19 May 2020].

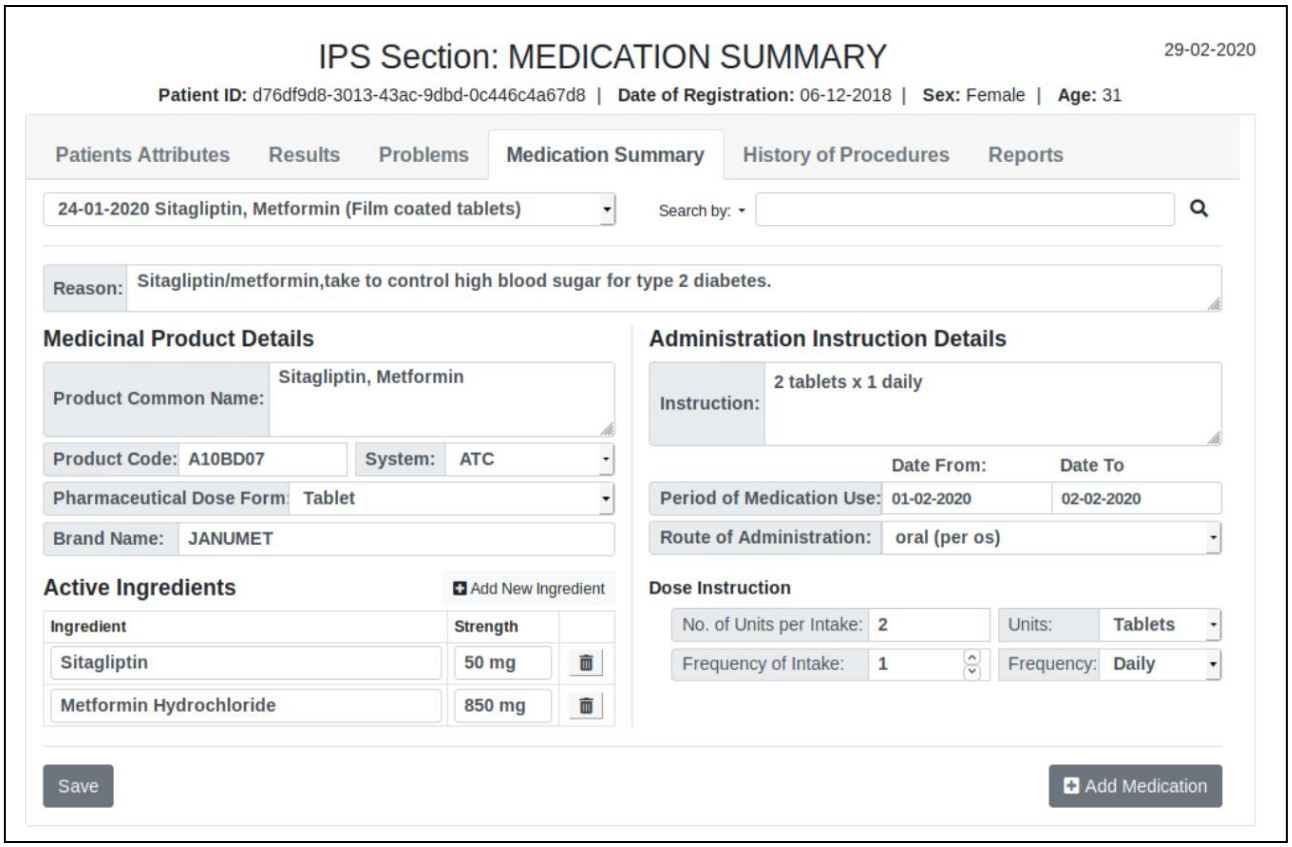


Figure 12. Snapshot of a client application adding a Medication to an openEHR archetype instance of the IPS Medication Summary Section.

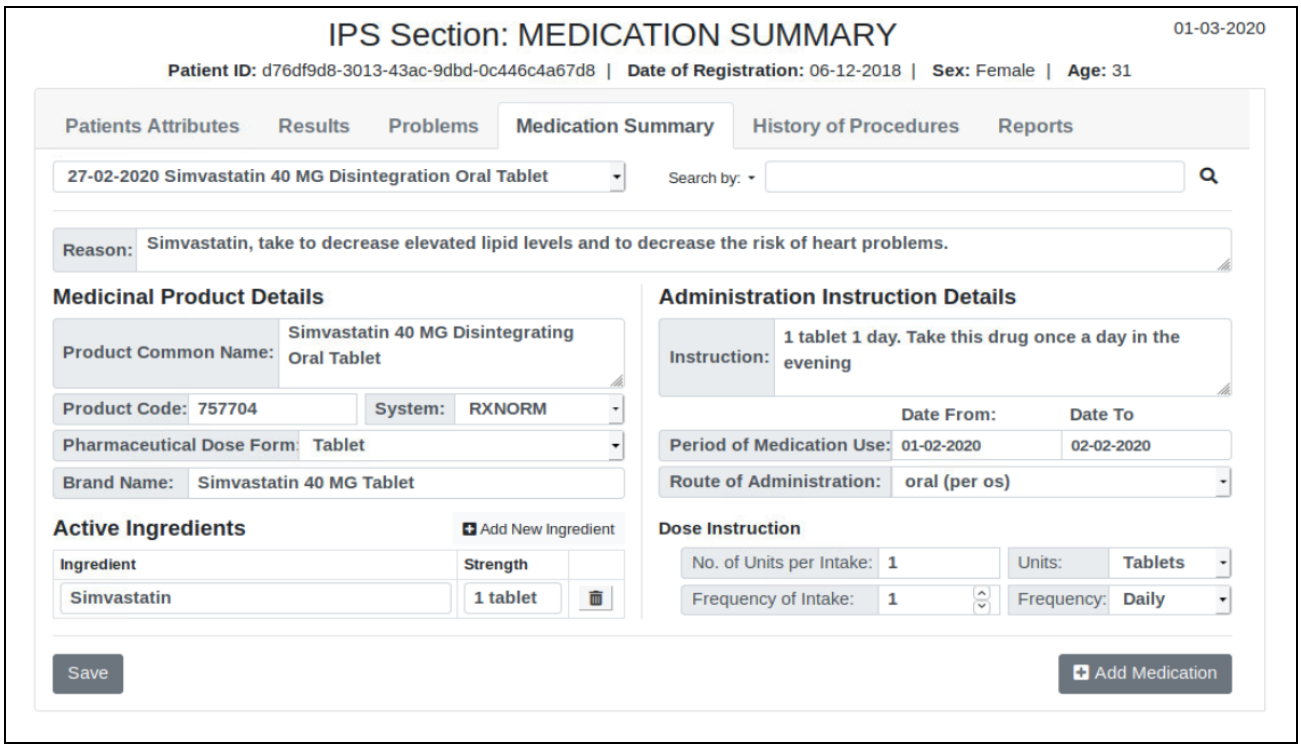


Figure 13. Snapshot of a client application updating an existing Medication in an openEHR archetype instance of the IPS Medication Summary Section.

Synthesis of Refinement Maps for Real-Time Object Code Verification

Eman M. Al-qtiemat*, Sudarshan K. Srinivasan*, Zeyad A. Al-Odat*, Sana Shuja†
*Electrical and Computer Engineering, North Dakota State University,
Fargo, ND, USA
†Department of Electrical Engineering, COMSATS University,
Islamabad, Pakistan
Emails: *eman.alqtiemat@ndsu.edu, *sudarshan.srinivasan@ndsu.edu, *zeyad.alodat@ndsu.edu,
†SanaShuja@comsats.edu.pk

***Abstract*—Refinement-based verification is a formal verification method, it is considered as a very scalable approach for dealing with low-level artifacts such as real-time object code verification. Two main obstacles prevent implementing the refinement-based verification; firstly, it requires formal specification in transition system form while most specifications are of informal or semi-formal form. To solve this issue, we already proposed synthesising procedures to transform both functional and timing requirements from natural language form into formal specifications, our approach was successfully applied on insulin pump safety requirements. Secondly, the verification process requires a construction of refinement map, which is a function maps implementation states (the artifact to be verified) onto specification states. Actually, constructing refinement maps often requires deep understanding and intuitions about the specification and implementation, it is shown very difficult to construct refinement maps manually. To go over this obstacle, the construction of refinement maps should be automated. As a first step toward the automation process, we manually developed refinement maps for various safety properties concerning the software control operation of insulin pumps. In addition, we identified possible generic templates for construction of refinement maps. To complement our previous work, this paper is built on refinement maps and refinement maps templates proposed previously to automate the construction of refinement maps. Synthesising procedures of refinement maps for functional and timing specifications are proposed. In addition, this paper shows more results of formal specifications and their suggested refinement map functions for timing requirements. Our work uses safety requirements of generic infusion pump model as heuristic data.**

Keywords–Formal verification; Synthesising of refinement maps; Formal specifications; Refinement-based verification.

I. Introduction

Software safety is one of the key challenges facing the design process [1] of safety-critical embedded systems such as medical devices [2]. For example, infusion pump (a medical device that delivers medication such as pain medication, insulin, cancer drugs etc., in controlled doses to patients intravenously) has 54 class 1 recalls related to software issued by the US Food and Drug Administration (FDA) [3]. Class 1 means that the use of the medical device can cause serious adverse health consequences or death.

Despite the fact that testing is the dominant verification technique currently used in commercial design cycles [4], testing can only show the presence of faults, but it never proves their absence [5]. Alternate verification processes should be applied to the software design in conjunction with testing to assure system correctness and reliability. Formal verification can address testing limitations by providing proofs of correctness for software safety. Intel [6], Microsoft [7] and [8], and Airbus [9] have successfully applied formal verification processes.

Refinement-based verification [10] is a formal verification technique that has been demonstrated to be effective for verification of software correctness at the object code level [11]. To apply refinement-based verification, software requirements should be expressed as a formal model. Previously, we have proposed a novel approach to synthesize formal specifications from natural language requirements [12], and in a later work, we have also addressed timing requirements and specifications [13].

Our verification approach is based on the theory of Well-Founded Equivalence Bisimulation (WEB) refinement [10]. In the context of WEB refinement, both the implementation and specification are treated as Transition Systems (TSs). If every behavior of the implementation is matched by a behavior of the specification and vice versa, then the implementation behaves correctly as prescribed by the specification. However, this is not easy to check in practice as the implementation TS and specification TS can look very different. The specification states obtained from the software requirements are marked with atomic propositions (predicates that are true or false in a given state). The implementation states are states of the microcontroller that the object code program modifies. As such, the microcontroller states includes registers, flags, and memory. The various possible values that these components can have during the execution of the object code program gives rise to the many millions of states of the implementation. To overcome this difference, WEB refinement uses the concept of a refinement map, which is a function that provided an implementation state, gives the corresponding specification state. Historically, one of the reasons that refinement-based verification is much less explored than other formal verification paradigms such as model checking is that the construction of refinement maps often requires deep understanding and intuitions about the specification and implementation [14]. However, once a refinement map is constructed, the benefit is that refinement-based verification is a very scalable approach for dealing with low-level artifacts such as real-time object code verification. In our previous paper [1], we have build refinement maps corresponding to formal specifications related to infusion pump safety and we have also proposed three possible generic refinement map templates, which is the first step toward automating the construction of refinement maps. This paper is based on our previous work, we propose synthesising procedures of refinement maps for both functional and timing requirements, the new procedure allow an expert

user intervention to assure the correctness of the system. The remainder of this paper is organized as follows. Section II summarizes background information. Section III details related work. Section IV describes the refinement maps and refinement map templates. Section V shows the proposed synthesis of refinement maps for system requirement. Conclusions and direction for future work are noted in Section VI.

II. BACKGROUND

This section explores the definition of transition systems, the definition of refinement-based verification, and the synthesis of formal specifications as key terms related to our work.

A. Transition Systems

As stated earlier, transition systems (TSs) are used to model both specification and implementation in refinement-based verification. TSs are defined below.

Definition 1: A TS $M = \langle S, R, L \rangle$ is a three tuple in which S denotes the set of states, $R \subseteq SXS$ is the transition relation that provides the transition between states, and L is a labeling function that describes what is visible at each state.

States are marked with Atomic Propositions (APs), which are predicates that are true or false in each state. The labeling function maps states to the APs that are true in every state. An example TS is shown in Figure 1. Here S = {S1, S2, S3, S4}, R = {(S1, S2), (S2, S4), (S4, S3), (S3, S4), (S3, S2), (S1, S3)} and, $L(S2)$ represents the atomic propositions that are true for the S2 state.

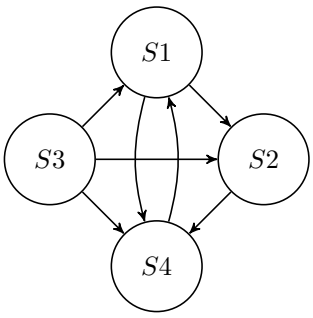


Figure 1. An example of a transition system (TS).

B. Timing Transition Systems

Some applications have requirements with timing conditions on the state's transitions called as timing requirements. Timing requirements explain the system behaviour under some timing constraints. Timing constraints are very important especially if we deal with a critical real time systems. As mentioned in the previous section (Section II-A), transition systems are used to represent the implementation and specification in refinement-based verification, however they do not contain timing requirements. Hence, in the verification of real time systems that contain timing constraints, timing transition systems (TTSs) [11] are used to represent the implementation and specification.

Definition 2: A TTS $M_t = \langle S, R_t, L \rangle$ is a three tuple in which S denotes the set of states and L is a labeling function that describes what is visible at each state. The state transition R_t has the form of $\langle x, y, l_t, u_t \rangle$ where $x, y \in S$ and $l_t, u_t \in N$ represents the lower and upper bounds as the timing condition for the transition.

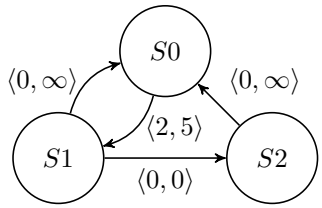


Figure 2. An example of a timing transition system (TTS).

Figure 2 shows an example of a timing transition system that consists of three states { S0, S1, S2 }, for instance; if the system is in state $S0$ it can go to state $S1$ only between 2 and 5 units of time, while going from $S1$ to $S2$ the time is zero meaning that it should happen immediately, going from $S2$ to $S1$ the time is zero to infinity which means that it can happen any point of time, and so on.

C. Refinement-Based Verification

Our verification approach is based on the theory of Well-Founded Equivalence Bisimulation (WEB) refinement. A detailed description of this theory can be found in [10]. Here, we give a very high-level overview of the key concepts. WEB refinement provides a notion of correctness that can be used to check an implementation TS against a specification TS. In the context of WEB refinement, both the implementation and specification are treated as TSs. The implementation behaves correctly as given by the specification, if every behavior of the implementation is matched by a behavior of the specification and vice versa. However, this is not easy to check in practice. Implementation TS and specification TS look quite different. The implementation states are states of the microcontroller that the object code program modifies. As such the microcontroller includes many registers, flags, and memory. The various possible values that these components can have during the execution of the object code program gives rise to the many millions of states of the implementation. To overcome this difference, WEB refinement uses the concept of a refinement map, which is a function that given an implementation states, tells you what is the corresponding specification state. Once a refinement map is constructed, WEB refinement verification proceeds as follows. The idea is to look at each implementation transition. Consider an implementation transition say (w, v), where both w and v are implementation states. To be correct, the transition has two options. The first option is that this implementation transition should match a specification transition, i.e., $r(w) = s$ and $r(v) = u$, where $r()$ is the refinement map, s and u are specification states, and (s, u) is a transition of the specification. This first option is called a non-stuttering implementation transition. The second option is that $r(w) = r(v) = s$, i.e., both w and v match to the same specification state. The second option is called a stuttering implementation transition. There are a few more checks to be performed. The very nice property of WEB refinement is that it is sufficient to reason about single steps of the implementation and specification to check for correctness and find bugs. This property makes WEB refinement very applicable to deal with the complexity of object code.

D. Synthesis of Functional Formal Specifications

Our approach for development and study of refinement maps is based on the formal TS specifications. We developed a previous approach to transform functional requirements into formal specifications [11]. Since this work is closely tied to the prior work, we briefly review it here. Figure 3 summarizes the transformation procedure, the main steps are explained as follows: functional requirement is fed as an input, an English parser called Enju is used to get the parse tree the requirement. The first step of computing the TSs is to apply Atomic Proposition Extraction Rule (APER) extract the APs from the requirements. We developed three Atomic Proposition Extraction Rules (APERs) that work on the parse tree of the requirement to get an initial list of APs. The resulting list is subjected to an expert user check (User Input), where the APs might be appended, eliminated or revised based on the expert users domain knowledge. A high-level procedure for specification transition system synthesis has been proposed to compute the states and transitions using the resulting list of APs under expert user supervision. Finally, the transition system is created using the resulting list of states and transitions. The output of the procedure is a formal specification TS.

E. Synthesis of Timing Formal Specifications

Some system requirements have timing constraints which are called timing requirements. We proposed a previous approach to work on transforming timing requirements into formal specifications [13]. Figure 4 shows the main steps of the synthesising procedure. A brief description of this approach is explained as follows: Timing requirement is fed as an input of the procedure. As in the previous procedure, Enju parser is used to get the parse tree that corresponds to the entered requirement. An Atomic Proposition and Timing Constrains Extraction Rule (APTCER) is applied on the resulted parse tree to get an initial list of APs and Timing Constrains (TCs). APs and TCs are paired together and they are considered the base of a TTS. Then, a set of states are defined based on the resulting list of APs. Transitions are applied between every two states. TCs are assigned to the transitions. This procedure allows input from domain expert as shown in Figure 4. Finally, the TTS is created. The output of this procedure is a formal specification TTS.

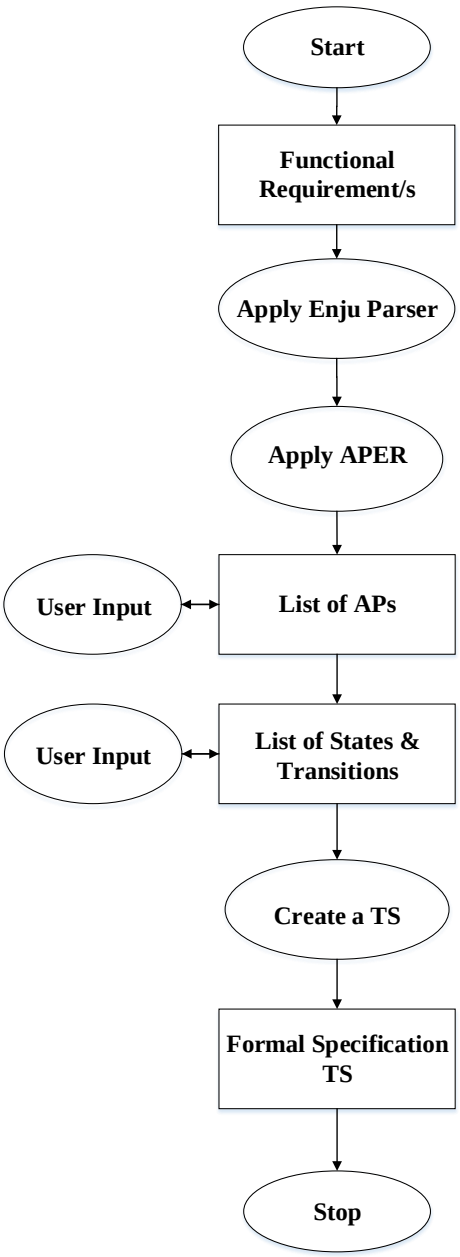


Figure 3. Formal Model synthesis procedure for Functional Requirements.

III. Related Work

This section summarizes a few works on applying refinement processes to get more concrete specifications and refinement-based verification. None of these works are applied to insulin pump formal specifications as our work. To the best of our knowledge, these are the most related state of art in this area of study.

Klein *et al.* [15] introduced a new technique called State Transition Diagrams (STD). It is a graphical specification technique that provides refinement rules, each rule defines an implementation relation on STD specification. The proposed approach was applied to the feature interaction problem. The refinement relation was utilized to add a feature or to define the notion of conflicting features.

Rabiah *et al.* [16] developed a reliable autonomous robot system by addressing A* path planning algorithm reliability issue. A refinement process was used to capture more concrete specifications by transforming High-Level specification into equivalent executable program. Traditional mathematical concepts were used to capture formal descriptions. Then, Z specification language was employed to transform mathematical description to Z schemas to get formal specifications. Z formal refinement theory was used to obtain the implementation specification.

Spichkova [17] proposed a refinement-based verification scheme for interactive real time systems. The proposed work solves the mistakes that rise from the specification problems by integrating the formal specifications with the verification system. The proposed scheme translates the specifications to a higher-order logic, and then uses the theorem prover (Isabelle) to prove the specifications. Using the refinement-based verification, this scheme validates the refinement relations between

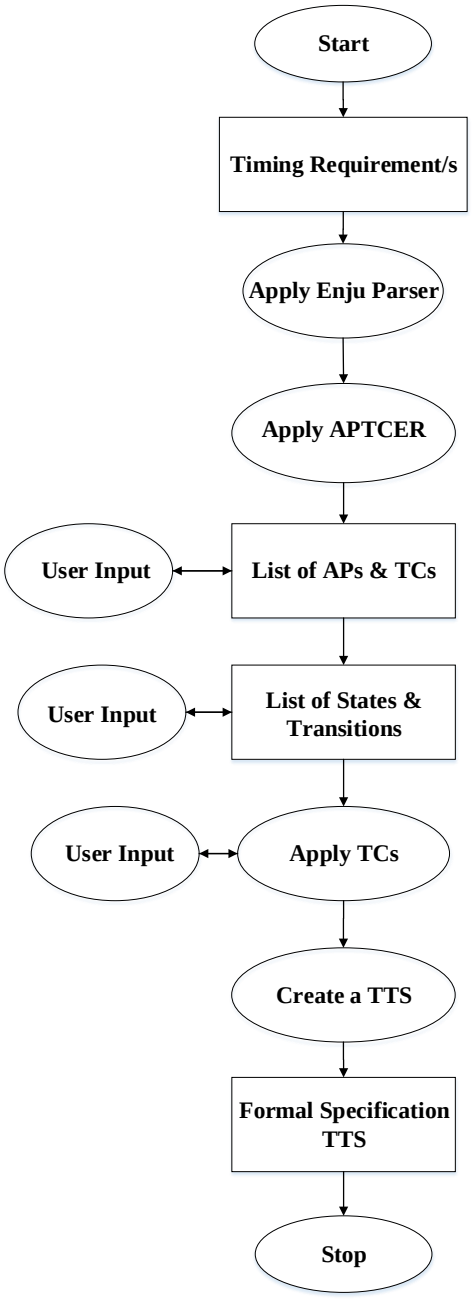


Figure 4. Formal Model synthesis procedure for Timing Requirements.

two different systems. The proposed design was tested and verified using a case study of electronic data transmission gateway.

A new approach that focuses on the refinement verification using state flow charts has been presented by Miyazawa *et al.* [18]. They proposed a refinement strategy that supports the sequential *C* implementations of the state flow charts. The proposed design benefited from the architectural features of model to allow a higher level of automation by retrieving the data relation in a calculation style and rendering the data into an automated system. The proposed design was tested and verified using Matlab Simulink SDK. Through the provided case study, the scheme was able to be scaled to different state charts problems.

Cimatti *et al.* proposed a contract-refinement scheme for embedded systems [19]. The contract-refinement provides interactive composition reasoning, step-wise refinement, and principled reuse refinements for components for the already designed or independently designed components. The proposed design addresses the problem of architectural decomposition of embedded systems based on the principles of temporal logic to generate a set of proof obligations. The testing and verification of the Wheel Braking System (WBS) case study show that the proposed design can detect the problems in the architectural design of the WBS.

Bibighaus [20] employed the Doubly Labeled Transition Systems (DLTS) to reason about possibilities security properties and refinement. This work was compared with three different security frameworks when applied to large class systems. The refinement framework in this work preserves and guarantees the liveness of the model by verifying the timing parameter of the model. The analysis results show that the proposed design preserves the security properties to a series of availability requirements.

A novel approach has been presented [21] to formally specify and analyze the certification process of Partitioning Operating Systems (POSs) by integrating refinement and ontology. An ontology of POSs was developed as an intermediate model between informal descriptions of ARINC 653 and the formal specification in Event-B. A semiautomatic translation has been implemented from the ontology and ARINC 653 into Event-B. Six hidden failures in ARINC were happened and fixed during the formal analysis in the Event-B specification. The existence of these errors has been validated in two open-source POSs: XtratuM and POK. The degree of automatic verification of the Event-B specification reached a higher level because of the ontology. By validation, they have also noticed some errors in open-source POSs. The proposed methodology has shown capability to to formalize and verify systems according to system's informal standards and requirements.

Human factors consider as the most obvious cause of failures especially when a human deals with critical systems such as nuclear and medical systems. A new methodology for developing a Human-Machine Interface (HMI) has been proposed [22], it uses a correct by construction approach. A HMI was developed independently using incremental refinement. Human interactions is dependent on testing, which can not guarantee the absence of failures. Formal method was used to assure the correctness of the human interactions. Even-B modeling language has been used to formalize the internal consistency with respect to the safety properties and events. This generic refinement strategy supports a development of the Model-View-Controller (MVC) architecture.

A specification development method and a generic security model were proposed based on refinement for ARINC Separation Kernels (KSs) [23]. A step-wise refinement framework was presented. Two levels of functional specification are developed by the refinement. Kernel initialization, inter-partition communication, two-level scheduling, and partition and process management were modeled. Isabelle/HOL theorem prover was used to carry out the formal specification and its security proofs. Mechanical check proofs were given to solve convert channels in separation kernels.

Fayolle *et al.* joined Algebraic State-Transition Diagrams (ASTD) with an Event-B specification for better understanding of the system behaviour [24]. They proposed an approach that works on incrementally refine the specification couplings, it takes the new refinement relations and consistency into consideration between data and control system specifications. This work had shown how to use two complementary languages for formal modeling, a railway CBTC-like case study were used. In addition, the principle of complementarity and consistency was explored between ASTD and B-like refinements. Separation between data and behavioural system's aspects were accomplished.

The issues of validating formal models were studied and executed using Event-B method [25]. Firstly, new techniques were created and discussed which allow model execution to be at all abstraction levels. To overcome barriers comes form non-deterministic features, users intervention such as modifying the model or providing ad hoc implementations were needed. Secondly, a new formal definition of the notion of fidelity was given, this definition assures specifying all the observable behaviors of the executable models by the non-deterministic models.

Many other papers discuss and analyze refinement concepts in the context of verifying concurrent objects. For example, Smith *et al.* in [26] provided formal link between trace refinement and linearizability, a comparison between these correctness conditions were explored. The main conclusion of this work is generally that trace refinement reveals linearizability, but linearizabilit does not reveals trace refinement. However, linearizability can reveal trace refinement but under specific conditions. Firstly, trace refinement can prove both safety and liveness properties, while linearizability can only prove safety properties. Secondly, the fact that trace refinement based on the identification of when the implementation operations are noticed to happen. They also studied these differences in the verification context of concurrent objects.

IV. Refinement Maps and Refinement Map Templates

Figures 5-11 show the formal specification TS for 8 insulin pump safety requirements, the figures also show the refinement map we developed corresponding to each specification. In this paper, formal specification TTS corresponding to 4 insulin pump timing requirements are added in Figures 12-15, we develop a refinement map for each specification as shown in the figures. The formal specifications TSs [12] and TTSs [13] were developed as part of our previous work in this area. As can be seen from the figures, each TS or TTS consists of a set of states and transitions between states. Also, each state is marked with the atomic propositions that are true. For TTSs in Figures 12-15, time bounds conditions are added on each transition. Our strategy for constructing the refinement maps is as follows. A specification state can be constructed from an implementation state by determining the APs that are true in the implementation state. If a specification has n APs, then we construct one predicate function for each AP. The predicate functions take the implementation state as input and output a predicate value that indicates if the AP is true in that state or not. Thus, the collection of such predicate functions is the refinement map.

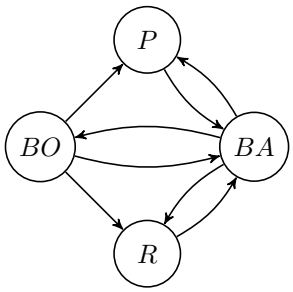
We next discuss the refinement map for the specification in Figure 5. The safety specification from [27] is as follows: "The pump shall suspend all active basal delivery and stop any active bolus during a pump prime or refill. It shall prohibit any insulin administration during the priming process and resume the suspended basal delivery, either a basal profile or a temporary basal, after the prime or refill is successfully completed." The APs corresponding to this safety requirement are (1) BO: active bolus delivery; (2) BA: active basal delivery; (3) P: priming process; and (4) R: refill process. The refinement map however has to account for what is happening in the implementation code and relate that to the atomic propositions.

The predicate function for BO uses several variables from the code including NB: Normal Bolus and EB: Extended Bolus as there are more than one type of Bolus dose supported by the system. So the AP BO should be true if there is a NB or an EB. NB is only a flag that indicates that a normal bolus should be in progress. The actual bolus itself will continue to occur as long as a counter that keeps track of the bolus has not reached its maximum value. Therefore, for example for a normal bolus, we use a conjunction of NB and the condition that the NB counter (NB_c) is less than its possible maximum value (NB_m). We use a similar strategy for the extended bolus as well. This refinement map template works for all processes similar to a Bolus dosage delivery, such as basal dosage delivery, priming process, and refill process. Therefore, we term this refinement map template as "process template." For the basal dosage (BA AP) a number of basal profiles (BPs) are possible that accounts for BP_1 thru BP_n. TB stands for temporary basal. As can be noted from Figures 6-15, the process template accounts for a large number of predicate functions corresponding to APs.

The second refinement map template is a simple one called the "projection template," which is used when the AP in the specification TS corresponds directly to a variable in the code. An example of the projection template can be found in Figure 6, where the User Reminder (UR) AP is mapped directly from a flag variable in the code that corresponds to the user reminder. A variation of this template is a boolean expression of Boolean variables in the code. An example of such an AP is the UIP AP in Figure 10.

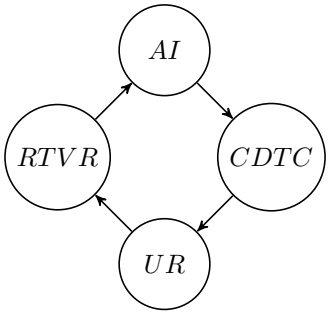
The third refinement map template is called the "value change template," which is used when the AP is true only when a value has changed. An example use of this template can be found in Figure 6 for the CDTC AP. CDTC corresponds to the change in drug type and concentration and is true when the drug type or concentration is changed. For the drug type change, DT is the variable that corresponds to the drug type. The question here is how to track that a value has changed. The idea is to use history variables. HDT is a history variable that corresponds to the history of the drug type, i.e., the value of the drug type in the previous cycle. If HDT is not equal to DT in a code state, then we know the drug type has changed. The inequality of HDT and DT is used to construct the predicate function. For all the safety requirements analyzed, these three refinement map templates cover all the APs.

For timed specifications, we next discuss the refinement map for the specification in Figure 13. The safety specification from [27] is as follows: "An air-in-line alarm shall be triggered within a maximum delay time of x seconds if air bubbles larger than y μL are detected, and all insulin administrations shall be stopped." The APs corresponding to this safety requirement



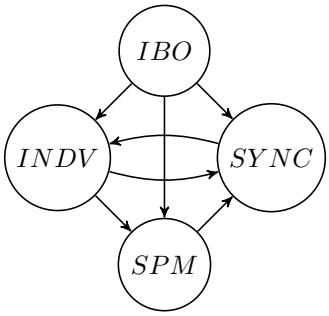
- **BO** = [NB $\wedge$ ($NB_c < NB_m$)] $\vee$ [EB $\wedge$ ($EB_c < EB_m$)]
- **P** = P $\wedge$ ($P_c < P_m$)
- **R** = R $\wedge$ ($R_c < R_m$)
- **BA** = [$BP_1 \wedge (BP_{1c} < BP_{1m})$] $\vee$ [$BP_2 \wedge (BP_{2c} < BP_{2m})$] $\vee \ldots \vee$ [$BP_n \wedge (BP_{nc} < BP_{nm})$] $\vee$ [TB $\wedge$ ($TB_c < TB_m$)]

Figure 5. A formal presentation of requirement 1.1.1 from [27] and the suggested refinement maps.



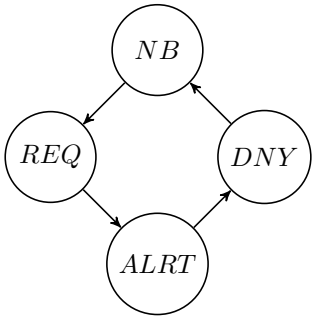
- **AI** = [$BP_1 \wedge (BP_{1c} < BP_{1m})$] $\vee$ [$BP_2 \wedge (BP_{2c} < BP_{2m})$] $\vee \ldots \vee$ [$BP_n \wedge (BP_{nc} < BP_{nm})$] $\vee$ [TB $\wedge$ ($TB_c < TB_m$)] $\vee$ [NB $\wedge$ ($NB_c < NB_m$)] $\vee$ [EB $\wedge$ ($EB_c < EB_m$)]
- **CDTC** = (DT $\neq$ HDT) $\wedge$ ($CDTC_c < CDTC_m$)
- **UR** = FLAG
- **RTVR** = (CRV $\neq$ HRV) $\wedge$ ($RTVR_c < RTVR_m$)

Figure 6. A formal presentation of requirement 1.1.3 from [27] and the suggested refinement maps.



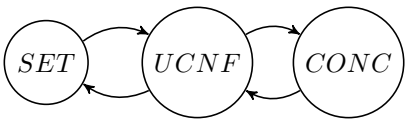
- **IBO** = [NB $\wedge$ ($NB_c < NB_m$)] $\vee$ [EB $\wedge$ ($EB_c < EB_m$)]
- **INDV** = [$BP_1 \wedge (BP_{1c} < BP_{1m})$] $\vee$ [$BP_2 \wedge (BP_{2c} < BP_{2m})$] $\vee \ldots \vee$ [$BP_n \wedge (BP_{nc} < BP_{nm})$] $\vee$ [TB $\wedge$ ($TB_c < TB_m$)] $\vee$ [NB $\wedge$ ($NB_c < NB_m$)] $\vee$ [EB $\wedge$ ($EB_c < EB_m$)]
- **SPM** = [P $\wedge$ ($P_c < P_m$)] $\vee$ [R $\wedge$ ($R_c < R_m$)]
- **SYNC** = INCAL $\wedge$ ($INCAL_c < INCAL_m$)

Figure 7. A formal presentation of requirement 1.8.2 and 1.8.5 from [27] and the suggested refinement maps.



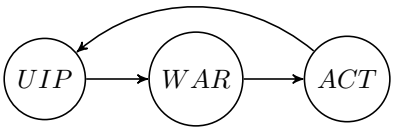
- **NB** = NB $\wedge$ ($NB_c < NB_m$)
- **REQ** = REQ-FLAG
- **ALRT** = ALRT-FLAG
- **DNY** = CALL-FUNCT

Figure 8. A formal presentation of requirement 1.3.5 from [27] and the suggested refinement maps.



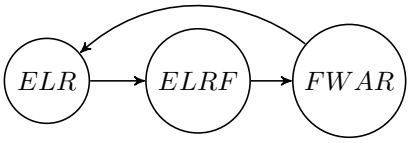
- **SET** = CLRS $\vee$ [CHNS $\wedge$ ($CHNS_c < CHNS_m$)] $\vee$ RESS
- **UCNF** = FLAG
- **CONC** = [SETT $\wedge$ ($SETT_c < SETT_m$)] $\vee$ [CHNC $\wedge$ ($CHNC_c < CHNC_m$)]

Figure 9. A formal presentation of requirement 2.2.2 and 2.2.3 from [27] and the suggested refinement maps.



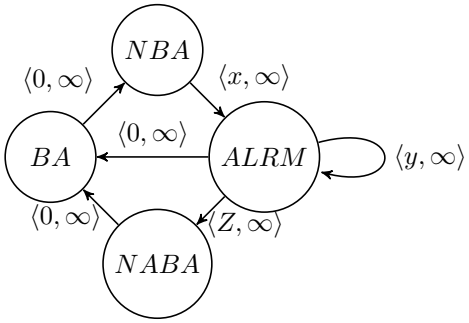
- **UIP** = BG $\vee$ TBG $\vee$ INCR $\vee$ CORF
- **WAR** = FLAG
- **ACT** = CNFI $\vee$ [CHNI $\wedge$ $(CHNI_c < CHNI_m)$]

Figure 10. A formal presentation of requirement 3.2.5 from [27] followed by the suggested refinement maps.



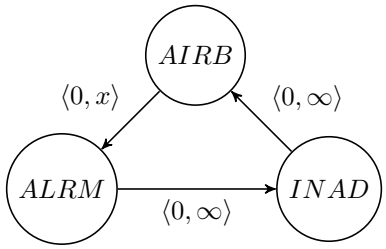
- **ELR** = [EL $\wedge$ $(EL_c < EL_m)$] $\vee$ [LR $\wedge$ $(LR_c < LR_m)$]
- **ELRF** = ELF $\vee$ LRF
- **FWAR** = FLAG

Figure 11. A formal presentation of requirement 3.2.7 from [27] followed by the suggested refinement maps.



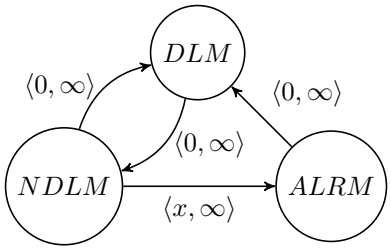
- **BA** = $[BP_1 \wedge (BP_{1c} < BP_{1m})] \vee [BP_2 \wedge (BP_{2c} < BP_{2m})] \vee \ldots \vee [BP_n \wedge (BP_{nc} < BP_{nm})]$ $\vee$ [TB $\wedge$ $(TB_c < TB_m)$]
- **NBA** = $\neg$ **BA**
- **ALRM** = ALRM-FLAG
- **NABA** = NBA $\wedge$ NA

Figure 12. A formal presentation of the timing requirement 1.2.8 from [27] and the suggested refinement maps.



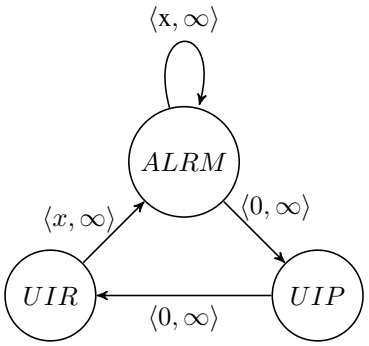
- **AIRB** = AB $>$ Y
- **ALRM** = ALRM-FLAG
- **INAD** = $[BP_1 \wedge (BP_{1c} < BP_{1m})] \vee [BP_2 \wedge (BP_{2c} < BP_{2m})] \vee \ldots \vee [BP_n \wedge (BP_{nc} < BP_{nm})]$ $\vee$ [TB $\wedge$ $(TB_c < TB_m)$] $\vee$ [NB $\wedge$ $(NB_c < NB_m)$] $\vee$ [EB $\wedge$ $(EB_c < EB_m)$]

Figure 13. A formal presentation of the timing requirement 1.6.1 from [27] followed by the suggested refinement maps.



- **DLM** = $[BP_1 \wedge (BP_{1c} < BP_{1m})] \vee [BP_2 \wedge (BP_{2c} < BP_{2m})] \vee \ldots \vee [BP_n \wedge (BP_{nc} < BP_{nm})]$ $\vee$ [TB $\wedge$ $(TB_c < TB_m)$] $\vee$ [NB $\wedge$ $(NB_c < NB_m)$] $\vee$ [EB $\wedge$ $(EB_c < EB_m)$]
- **NDLM** = $\neg$ **DLM**
- **ALRM** = ALRM-FLAG

Figure 14. A formal presentation of the timing requirement 1.8.4 from [27] followed by the suggested refinement maps.



- **ALRM** = ALRM-FLAG
- **UIR** = $RTD \vee RDC \vee RRDV \vee \ldots$
- **UIP** = $STD \vee SDC \vee RDV \vee \ldots$

Figure 15. A formal presentation of the timing requirement 2.2.1 from [27] followed by the suggested refinement maps.

are (1) AIRB: air bubbles; (2) ALRM: air-in-line alarm; (3) INAD: insulin administration. The predicate function for INAD uses several variables from the code including BPs; TB; NB; and EB as explained above. The AP INAD should be true if one of these variables is true and its counter variable is less than the maximum value. This AP is considered as an example use of the process template. For the ALRM AP, it is a simple example of the projection template, it should be true if its corresponding flag is true. The AIRB AP shows another variation of the value change template, which is depends on the changing value of the Air Bubbles (AB) variable. If the AB variable value is greater than Y (Y is a predefined value of the number of bubbles and it is based on the pump model), the AirB AP will be true. Table I gives the expansions for all the abbreviations used in Figures 6-15, so that the corresponding refinement maps can be comprehended by the reader.

V. Synthesis of Refinement Maps for System Requirements

This section explains new automation procedures for constructing refinement maps for both functional and timing system requirements from [27]. The first part of this work uses our previously proposed algorithms for synthesising formal specifications from natural language requirements [12] [13].

Procedure 1 shows the overall flow for computing the refinement map template for each AP in a functional requirement. A set of functional requirements in natural language form is fed as input. Three template lists are the output of the procedure, each list will contain a set of APs based on the heuristic data from the parsed trees belonging to each input requirement.

Three empty template lists are defined; projection template list (line 1), process template list (line 2), and value template list (line 3). A list for functional requirement's APs (AP_f-list) is initialized to null (line 4). Each requirement is input to the Synthesising Procedure for Functional Requirements (SPFR) (line 6) which comes up with formal specifications (explained in II-D). A function called Get_AP-list is used to obtain the resulting AP-list from the SPFR into the AP_f-list (line 7). A function called Get_Sub-tree is applied to each entry (AP_f) in the AP_f-list, this function returns the sub tree that corresponds to the AP_f from Enju parsed tree (line 9). A function called Head stores the head category of the sub tree in variable X (line 10). Check if X is of NX category, right child of X is PP, and left child of X is NX (line 11), then AP is added to the projection template list (line 12). Also if X is of VP category, right child of X is NP, and left child of X is VX (line 11), so AP is added to the projection template list as well. For the AP to be stored in the value change template (line 15), there are two cases; case 1: If X is of NX category, left child of X is VP, and right child is NX-COOD (line 14). Case 2: if X is of VP category, right child of X is CP, and left child of X is VX (line 14). If the sub tree of AP does not meet any of the previous mentioned conditions, then AP will be stored in the process template list (line 17). The procedure allows

Procedure 1 Procedure for synthesizing Refinement Maps for functional requirements

Require: Set of Functional Requirements
1: Projection-Temp-list ← ∅ ;
2: Process-Temp-list ← ∅ ;
3: Value-Temp-list ← ∅ ;
4: AP_f-list ← ∅ ;
5: **for each** $Req_f \in$ Functional Requirements **do**
6: Apply_SPFR(Req_f);
7: AP_f-list ← Get_AP-list(SPFR);
8: **for each** $AP_f \in AP_f$-list **do**
9: Sub-tree ← Get_Sub-tree (AP_f);
10: X = Head(Sub-tree);
11: **if** [(X = NX) ∧ (RightChild(X) = PP) ∧ (LeftChild(X) = NX)] ∨ [(X = VP) ∧ (RightChild(X) = NP) ∧ (LeftChild(X) = VX)] **then**
12: Projection-Temp-list ← Projection-Temp-list ∪ AP_f;
13: **else**
14: **if** [(X = NX) ∧ (LeftChild(X) = VP) ∧ (RightChild(X) = NX-COOD)] ∨ [(X = VP) ∧ (RightChild(X) = CP) ∧ (LeftChild(X) = VX)] **then**
15: Value-Temp-list ← Value-Temp-list ∪ AP_f;
16: **else**
17: Process-Temp-list ← Process-Temp-list ∪ AP_f;
18: Projection-Temp-list ← USR_IN(Projection-Temp-list);
19: Value-Temp-list ← USR_IN(Value-Temp-list);
20: Process-Temp-list ← USR_IN(Process-Temp-list);

TABLE I. LIST OF ABBREVIATIONS

Abbreviation	Meaning
AI	Active Infusion
CDTC	Change Drug Type and Concentration
DT	Data Type
HDT	Historical Data Type
UR	User Reminder
RTVR	Reservoir Time and Volume Recomputed
CRV	Current Reservoir Volume
HRV	Historical Reservoir Volume
IBO	Incomplete Bolus
INDV	Insulin Delivery
SPM	Suspension Mode
SYNC	Synchronization
INCAL	Insulin Calculations
REQ-FLAG	Request Flag
CALL-FUNCT	Call-Function for Calculation
SET	Settings
CLRS	Clear Settings
CHNS	Change Settings
RESS	Reset Settings
UCNF	User Confirmation
SETT	Setting the concentration
CHNC	Changing the Concentration
BG	Blood Glucose
TBG	Targeted Blood Glucose
INCR	Insulin to Carbohydrate ratio
CORF	Correction Factor
ACT	User Action
CNFI	Confirm Input
CHNI	Change Input
ELR	Event or Log Retrieving
EL	Event Logging
LR	Log Retrieving
ELRF	Event Logging or Logging Retrieving Failure
ELF	Event Logging Failure
LRF	Logging Retrieving Failure
ELF	Event Logging Failure
FWAR	Failure Warning
NBA	NO Basal delivery
NABA	No Alarm or Basal delivery
NA	No Alarm
AB	Air Bubbles
DLM	Delivery Mode
NDLM	Non-Delivery Mode
UIR	User Input Requested
RTD	Requested Time and Date
RDC	Requested Drug type and Concentration
RRDV	Requested Reloading Drug reservoir
UIP	User Input Provided
STD	Setting Time and Date
SDC	Setting Drug type and Concentration
RDV	Reloading Drug reservoir

Procedure 2 Procedure for synthesizing Refinement Maps for timing requirements

Require: Set of Timing Requirements
1: Projection-Temp-list $\leftarrow \emptyset$;
2: Process-Temp-list $\leftarrow \emptyset$;
3: Value-Temp-list $\leftarrow \emptyset$;
4: AP_t-list $\leftarrow \emptyset$;
5: **for each** $Req_t \in$ Timing Requirements **do**
6: Apply_SPTR(Req_t);
7: AP_t-list $\leftarrow$ Get_AP-list (SPTR);
8: **for each** $AP_t \in$AP$_t$-list **do**
9: Sub-tree $\leftarrow$ Get_Sub-tree (AP_t);
10: $X1$ = Head(Sub-tree);
11: **if** [($X1$ = NX) $\wedge$ (RightChild($X1$) = NX) $\wedge$ (LeftChild($X1$) = ADJ)]$\vee$ [($X1$ = NX) $\wedge$ (RightChild($X1$) = NX) $\wedge$ (LeftChild($X1$) = NP)] **then**
12: Projection-Temp-list $\leftarrow$ Projection-Temp-list $\cup$ AP_t;
13: **else**
14: **if** [($X1$ = NX) $\wedge$ (RightChild($X1$) = ADJ) $\wedge$ (LeftChild($X1$) = NX)] **then**
15: Value-Temp-list $\leftarrow$ Value-Temp-list $\cup$ AP_t;
16: **else**
17: Process-Temp-list $\leftarrow$ Process-Temp-list $\cup$ AP_t;
18: Projection-Temp-list $\leftarrow$ USR_IN(Projection-Temp-list);
19: Value-Temp-list $\leftarrow$ USR_IN(Value-Temp-list);
20: Process-Temp-list $\leftarrow$ USR_IN(Process-Temp-list);

expert user input to the final template lists (lines 18-20), the user can modify, delete, add or exchange APs from any list if any AP is classified in the wrong list. Procedure 2 shows the overall flow for computing the refinement map template for each AP in a timing requirement. A set of timing requirements in natural language form is fed as an input. Three templates lists are the output of the procedure as in procedure I, each list will contain a set of APs based on the heuristic data from the parsed trees belonging to each input requirement. Three empty template lists are defined; projection template list (line 1), process template list (line 2), and value template list (line 3). A list for APs (AP_t-list) is initialized to null (line 4). Each requirement is input to the Synthesising Procedure for Timing Requirements (SPTR) (line 6) which comes up with formal specifications (explained in Section II-E). A function called Get_AP-list is used to obtain the resulting AP-list from the SPTR into the AP_t-list (line 7). A function called Get_Sub-tree is applied to each entry (AP_t) in the AP_t-list, this function returns the sub tree that corresponds to the AP_t from Enju parsed tree (line 9). A function called Head stores the head category of the sub tree in variable $X1$ (line 10). Check if $X1$ is of NX category, right child of $X1$ is NX, and left child of $X1$ is ADJ (line 11), then AP is added to the projection template list (line 12). Also if $X1$ is of NX category, right child of $X1$ is NX, and left child of $X1$ is NP (line 11), so AP is added to the projection template list. For the AP to be stored in the value change template (line 15), there is only one case;

If $X1$ is of NX category, right child of $X1$ is of ADJ category and left child of $X1$ is also NX (line 14). If the sub tree of AP does not meet any of the previous mentioned conditions, then it will be stored in the process template list (line 17). As in procedure I, this procedure allows expert user input to the final template lists (lines 18-20), the user can modify, delete, add or exchange APs from any list if any AP is classified in the wrong list.

VI. Conclusion and Future Work

In this paper, we have developed a process for synthesizing refinement maps. Heuristics have been developed based on the output of the Enju parser to select a refinement map template for each atomic proposition. The key ideas of our approach are the following. Firstly, the system requirement is fed as an input. Secondly, the previously proposed synthesising procedure of formal specifications is applied on the input requirement. Finally, the heuristic data from the requirement's parsed tree is used to select the suitable refinement map template. The refinement map template is either the process template, the projection template or changing value template. For future, our work can be applied on any critical device that has safety requirements, and more generic refinement map templates can be identified. In addition, the level of automation can be increased by improving the synthesis procedure.

Acknowledgment

This publication was funded by a grant from the United States Government and the generous support of the American people through the United States Department of State and the United States Agency for International Development (USAID) under the Pakistan - U.S. Science & Technology Cooperation Program. The contents do not necessarily reflect the views of the United States Government.

References

[1] E. M. Al-Qtiemat, S. K. Srinivasan, Z. A. Al-Odat, and S. Shuja, "Refinement maps for insulin pump control software safety verification," in The Eleventh International Conference on Advances in System Testing and Validation Lifecycle VALID 2019. IARIA.

[2] B. Fei, W. S. Ng, S. Chauhan, and C. K. Kwoh, "The safety issues of medical robotics," Reliability Engineering & System Safety, vol. 73, no. 2, 2001, pp. 183–192.

[3] FDA, "List of Device Recalls, U.S. Food and Drug Administration (FDA)," 2018, last accessed: 2019-10-11. [Online]. Available: https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfRES/res.cfm

[4] S. Quadri and S. U. Farooq, "Software testing-goals, principles, and limitations," International Journal of Computer Applications, vol. 6, no. 9, 2010, pp. 7–10.

[5] E. Miller and W. E. Howden, Tutorial, software testing & validation techniques. IEEE Computer Society Press, 1981.

[6] R. Kaivola et al., "Replacing testing with formal verification in intel coretm i7 processor execution engine validation," in Computer Aided Verification, 21st International Conference, CAV, Grenoble, France, June 26 - July 2, 2009. Proceedings, pp. 414–429. [Online]. Available: https://doi.org/10.1007/978-3-642-02658-4_32

[7] T. Ball, B. Cook, V. Levin, and S. K. Rajamani, "SLAM and static driver verifier: Technology transfer of formal methods inside microsoft," in Integrated Formal Methods, 4th International Conference, IFM, Canterbury, UK, April 4-7, 2004, Proceedings, pp. 1–20. [Online]. Available: https://doi.org/10.1007/978-3-540-24756-2_1

[8] K. Bhargavan et al., "Formal verification of smart contracts: Short paper," in Proceedings of the 2016 ACM Workshop on Programming Languages and Analysis for Security. ACM, 2016, pp. 91–96.

[9] D. Delmas, E. Goubault, S. Putot, J. Souyris, K. Tekkal, and F. Védrine, "Towards an industrial use of fluctuat on safety-critical avionics software," in International Workshop on Formal Methods for Industrial Critical Systems. Springer, 2009, pp. 53–69.

[10] P. Manolios, "Mechanical verification of reactive systems," PhD thesis, University of Texas at Austin, August 2001, last accessed: 2019-10-04. [Online]. Available: http://www.ccs.neu.edu/home/pete/research/phd-dissertation.html

[11] M. A. L. Dubasi, S. K. Srinivasan, and V. Wijayasekara, "Timed refinement for verification of real-time object code programs," in Working Conference on Verified Software: Theories, Tools, and Experiments. Springer, 2014, pp. 252–269.

[12] E. M. Al-qtiemat, S. K. Srinivasan, M. A. L. Dubasi, and S. Shuja, "A methodology for synthesizing formal specification models from requirements for refinement-based object code verification," in The Third International Conference on Cyber-Technologies and Cyber-Systems. IARIA, 2018, pp. 94–101.

[13] E. M. Al-Qtiemat, S. K. Srinivasan, Z. A. Al-Odat, and S. Shuja, "Synthesis of Formal Specifications From Requirements for Refinement-based Real Time Object Code Verification," International Journal on Advances in Internet Technology, vol. 12, Aug 2019, pp. 95–107.

[14] M. Abadi and L. Lamport, "The existence of refinement mappings," Theoretical Computer Science, vol. 82, no. 2, 1991, pp. 253–284.

[15] C. Klein, C. Prehofer, and B. Rumpe, "Feature specification and refinement with state transition diagrams," arXiv preprint arXiv:1409.7232, 2014.

[16] E. Rabiah and B. Belkhouche, "Formal specification, refinement, and implementation of path planning," in 12th International Conference on Innovations in Information Technology (IIT). IEEE, 2016, pp. 1–6.

[17] M. Spichkova, "Refinement-based verification of interactive real-time systems," Electronic Notes in Theoretical Computer Science, vol. 214, 2008, pp. 131–157.

[18] A. Miyazawa and A. Cavalcanti, "Refinement-based verification of sequential implementations of stateflow charts," arXiv preprint arXiv:1106.4094, 2011.

[19] A. Cimatti and S. Tonetta, "Contracts-refinement proof system for component-based embedded systems," Science of computer programming, vol. 97, 2015, pp. 333–348.

[20] D. L. Bibighaus, "Applying doubly labeled transition systems to the refinement paradox," Naval Postgraduate School Monterey CA, Tech. Rep., 2005.

[21] Y. Zhao, D. Sanán, F. Zhang, and Y. Liu, "Formal specification and analysis of partitioning operating systems by integrating ontology and refinement," IEEE Transactions on Industrial Informatics, vol. 12, no. 4, 2016, pp. 1321–1331.

[22] R. Geniet and N. K. Singh, "Refinement based formal development of human-machine interface," in Federation of International Conferences on Software Technologies: Applications and Foundations. Springer, 2018, pp. 240–256.

[23] Y. Zhao, D. Sanán, F. Zhang, and Y. Liu, "Refinement-based specification and security analysis of separation kernels," IEEE Transactions on Dependable and Secure Computing, vol. 16, no. 1, 2017, pp. 127–141.

[24] T. Fayolle, M. Frappier, R. Laleau, and F. Gervais, "Formal refinement of extended state machines," arXiv preprint arXiv:1606.02016, 2016.

[25] A. Mashkoor, F. Yang, and J.-P. Jacquot, "Refinement-based validation of event-b specifications," Software & Systems Modeling, vol. 16, no. 3, 2017, pp. 789–808.

[26] G. Smith and K. Winter, "Relating trace refinement and linearizability," Formal Aspects of Computing, vol. 29, no. 6, 2017, pp. 935–950.

[27] Y. Zhang, R. Jetley, P. L. Jones, and A. Ray, "Generic safety requirements for developing safe insulin pump software," Journal of diabetes science and technology, vol. 5, no. 6, 2011, pp. 1403–1419.