Digital Personalized Health and Medicine
L.B. Pape-Haugaard et al. (Eds.)

doi:10.3233/SHTI200181

Ontology-Based and Architecture-Based Method for the Development of Interoperable Care Systems for Type 2 Diabetes Mellitus

Gustavo A. URIBE[a,b,c,1], Bernd BLOBEL[d,e,f], Diego M. LOPEZ[c], Stefan SCHULZ[g], and Alonso A. RUIZ[h]
[a] *The European Organization for Nuclear Research, Geneva, Switzerland*
[b] *Experimental High Energy Physics, University Antonio Nariño, Bogotá, Colombia*
[c] *Telematics Engineering Research Group, University of Cauca, Popayán, Colombia*
[d] *Medical Faculty, University of Regensburg, Regensburg, Germany*
[e] *eHealth Competence Center Bavaria, Regensburg, Germany*
[f] *First Medical Faculty, Charles University Prague, Prague, Czech Republic*
[g] *Institute of Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria*
[h] *Medical Faculty, University of Cauca, Popayán, Colombia*

Abstract. Processes like the care of type 2 diabetes mellitus patients require support by information systems considering the heterogeneity of the actors from different domains involved, enabling harmonization and integration of their specific methodologies and knowledge representation approaches towards interdisciplinary cooperation. Currently, the development of systems starts from the simplified information world, ignoring the aforementioned heterogeneity and specificity of real-world processes. This paper aims to demonstrate the feasibility of developing an adaptive, interoperable and intelligent system that supports the major aspects of type 2 diabetes mellitus care based on the Generic Component Model as formal methodology for modelling universal systems. The result is a deployable solution based on a formal representation of the diabetes care system, its objectives, and the intended business process. The implemented system enables reasoning over the data, inferring medical diagnosis. The effectiveness of the inference was evaluated, obtaining an F-measure of 0.89. The methods presented in this paper helps to build high quality models based on computation-independent aspects, which enable the construction of knowledge-based adaptive, intelligent and interoperable eHealth systems.

Keywords. Diabetes mellitus, Health information management, Ontology-based Systems, Systems architecture, Knowledge-based systems

1. Introduction

Currently, numerous software applications have been developed and implemented to support health care services. Few of them aim at achieving data interoperability, mapping

[1] Gustavo A. Uribe, ATLAS experiment, CERN, Espl. des Particules 1, 1211 Meyrin, Switzerland; E-mail: gustavo.uribe@cern.ch.

the structure of their data to different standards. Even fewer systems consider the concepts and contexts of the data represented, i.e., the knowledge of the actors and the interoperability problems associated with the heterogeneity of this knowledge. Therefore, most of the systems have difficulties to "understand" the data and to adapt them consistently and correctly to the different stakeholders. Improving health care services by interpreting data to perform conclusions and actions requires a specific level of system intelligence, especially needed in low and middle income countries where skilled human resources are limited.

The deployed methodology, i.e. the architectural model and framework, for designing and implementing a type 2 diabetes mellitus (T2DM) software system is based on the Generic Component Model (GCM) [1]. GCM allows the representation of any system architecture and the analysis and design of information systems. The GCM has been explained, discussed, used and compare in several papers, e.g. [1-4]. However, most of those papers focus on the principles GCM, without evaluating concrete software implementations. The proposed method allows the harmonization of the diversity of terms and concepts involved in health services by abstracting them. Thus, interoperability, collaboration and decision-making between health systems actors are demonstrated for a T2DM care scenario in the Colombian context. The paper's main objective is to demonstrate the feasibility of developing an adaptive, interoperable and intelligent T2DM care system based on a formal methodology for modelling of systems and their ontological representation. The T2DM scenario was considered because of its high prevalence, 8.5% in 2014 [5], and its complex management due the required collaboration of many health professionals from different domains.

2. Methods

The methodology followed in this work is based on the GCM framework for analysis, design and implementation of the systems, providing a generic abstract architecture for each domain-specific aspect of the system instantiated by domain-specific ontologies. The obtained model addresses three domains, viz. the medical, policy and resource domain. The upper-level ontology BioTopLite [6] was deployed to harmonize the different domains' concepts and knowledge. It provides high compatibility with the top-level ontology BFO [7], and it additionally considers some relevant and general aspects of the biological domain.

Ontology languages such as the Web Ontology Language (OWL) were used to represent architectural and static aspects of the system such as the generic system description and the description of the glycaemic control use case. Functional or dynamic aspects of the system are modelled by the Business Process Modeling Notation (BPMN) [8]. The business process model describes the expected behaviour of a business system and allows controlling, monitoring or implementing the ongoing processes.

Policies and rules governing the system for the glycaemic control use case were formally described using the SPARQL Inference Notation (SPIN) [9]. SPIN is a general purpose, standardized, and broadly accepted rule language that allows the definition of rules related to ontology classes. Now, the computation-independent T2DM care system was completely modelled. Based on this description, the software development process was performed according to the ISO 10746 RM-ODP [10] and its viewpoints as defined in the GCM. The output of this process is the specification of the information technology (IT) solution and, finally, the running application system.

3. Results

Generic and specialized architectural models of the system have been explained in [11-13]. For each GCM viewpoint, the models are adapted according to some inputs required for the development process [1, 13]. The Business View corresponds to the architectural models expressed using OWL and SPIN, and semi-formal models of the behaviour represented using BPMN. The Enterprise View defines the roles, activities and policies statements of the specified system [10]. The actors' roles in the system can be classified into: the categories of health organization staff, the self-care actor, the organizational administrator, and the resource chief. The roles have been derived from the International Standard Classification of Occupations [14]. The Information View represents in a platform-independent way the semantics of information [10]. It is defined in the ontology, but the datatypes of the information are not. In our approach, the entities representing data are *btl2:InformationObject* individuals. These entities are the only ones that use datatypes in a computational sense. The Computation View describes the functional decomposition of the system [10]. A functional decomposition was performed according to the general information cycle for any collaboration [13]. Five generic components has been defined: the data storage, the data interpreter (processing the data according to the formalized knowledge), the data mapper (maps the information to the knowledge of specific actor), the planner (decide the plan to be followed), the execution controller (assist the actors in the plan execution), and the execution listener (obtains new data after the plan execution). Additionally, some support components an execution engine, the reasoner, and the user interfaces are defined. The Engineering View describes the platform-specific infrastructure that supports the computational view. In our implementation, the system was deployed in four devices, the self-care device, the health professional device, a node including most of the computational components, and a node including the data storage. All are connected by Internet Protocol version 4 and HTTP. Finally, the Technology View describes the implementation of the system in terms of a configuration of technology objects representing the hardware and software components of the implementation [10]. In this view, the technologies used to implement the functional components are selected. For the execution engine functional component, the CamundaBPMN version 7.3 was selected. For the reasoner component the SPIN Rule Reasoner in its version 1.4 was deployed, and for the data storage the Virtuoso Openlink Server version 6.1 was used. The other functional components have been implemented using the Spring Framework version 4.1.7, available as open-source projects [15].

For evaluating the proposed system, a quantitative formal experiment was performed. Effectiveness is quantified using precision, recall and F-measure metrics. The hypothesis of the experiment is that the F-measure of the system's diagnosis is higher than 0.71, it is the average of the algorithms C4.5 and CART evaluated for the diagnosis of diabetes [16]. The specialist provided 20 anonymized medical records including its observations and decisions made for these patients. The medical records were structured according with the proposed information model, and then 1200 medical observations were obtained [15]. The prototype developed operates as an expert system for glycaemic control. after entering the 20 medical records, the system provides diagnostic suggestions. Table 1 compares the outcomes for one medical record. Some diagnoses corresponding to non-diabetic complications (e.g. chondromalacia of patella, mild malnutrition) were not included in the system. Other diagnoses have not been asserted due to the difficulty to infer them using the aforementioned rules, such as no chronic complications, uncomplicated diverticular disease colon, or probable primary hypothyroidism. Some correct states about the patient have been added by the designed system (bolded in Table 1). These diagnoses were considered irrelevant by the specialist because they are not included in ICD10, which is used to classify the relevant medical diagnoses in Colombia. The mean F-measure obtained was 0.89, with a minimum value of 0.57 and a maximum value of 1. The F-measure has a standard deviation of 0.13. The mean precision is 1 and recall is 0.82. The significance of the results was evaluated with a one-sample T-test. The one-sample T-test resulted in an F-measure significantly higher than the threshold value (0.71) with a value of $p=0.00$.

Table 1. Diagnoses by physicians versus diagnoses by the system (underlined and bolded the differences)

Diagnosis by Physician	Diagnosis by the System
Type 2 diabetes mellitus	Type 2 diabetes mellitus
Peripheral diabetic neuropathy	Peripheral diabetic neuropathy
Overweight	Overweight
Metabolic syndrome	Metabolic syndrome
Hypertriglyceridemia	Hypertriglyceridemia
Raised fasting plasma glucose	Raised fasting plasma glucose
Scleral and hypertensive cardiopathy	**South Asian central obesity**
Congestive heart failure stage II – C	**Decreased ankle reflex**
Coronary artery disease	**Medical alert – hyperglycemia**
	Hypoesthesia
	Medical alert – hypertension

4. Discussion

The correct and formal description of domains and contexts through rules allows the adaptability of the system. For example, the presented system was developed using a description of the context of the Colombian health system, but it can be adapted to other health systems by the definition of their specific context. Furthermore, the software system is able to adapt to the physician's preferred language and methodologies including internal guidelines, measurement units, etc.

Contrary to the presented proposal, traditionally, the development team models the system by abstracting the types of information generated and shared during the business process. The resulting models are semi-formal descriptions of the system and are not intended to formally describe the system domains' knowledge, thereby preventing knowledge sharing and management including the deployment of best-of-breed expertise. Therefore, at least the following problems arise in traditional development processes:

1. The models are highly dependent on the knowledge of the development team. Different development teams provide heterogeneous models without a clear way of harmonization.
2. The models ignore essential concepts and knowledge of the business domain, as domain experts are usually not part of the team.
3. The models cannot guarantee correct inferences using logic rules.
4. Most parts of the models are specific for the correspondent business process, limiting the re-usability of components and reducing the chance of interoperability.
5. There is no a clear separation between the business domain description and the description of the information objects. That makes interoperability between information models difficult.

One limitation of the presented proposal is the medically relevant diagnoses of absence conditions, e.g. the diagnoses “No chronic complications” and “Uncomplicated diverticular disease colon”.

5. Conclusions

This paper presents the development of a software system and its evaluation process, starting from an abstract description of the real system considering computation-independent aspects up to a concrete ontology-based web application. This application follows medical guidelines, includes the data of the patients and suggests a precise diagnosis.

The paper testifies the feasibility to develop an adaptive, interoperable and intelligent T2DM Care System based on the GCM methodology and framework. The resultant software is available as open source in [15].

References

[1] Blobel, Bernd, "Challenges and Solutions for Designing and Managing pHealth Ecosystems", Frontiers in medicine, vol. 6, article 83, 2019.

[2] Blobel, B. and Oemig, F., “Solving the modeling dilemma as a foundation for interoperability”, European Journal for Biomedical Informatics, vol. 14, no. 3, pp.3-12, 2018.

[3] Blobel, Bernd, and Frank Oemig. "The importance of architectures for interoperability", *Stud Health Technol Inform*, vol. 211, pp. 18–56, 2015.

[4] Health Level 7 International, “HL7 Version 3 Standard: Security and Privacy Ontology, Release 1.0,” *Ann Arbor: HL7 International*, 2013.

[5] “Diabetes”, *World Health Organization*, 2015. [Online]. Available: http://www.who.int/news-room/fact-sheets/detail/diabetes. [Accessed: 03-Aug-2018].

[6] S. Schulz and M. Boeker, “BioTopLite: An Upper Level Ontology for the Life Sciences. Evolution, Design and Application”, presented at the Workshop on Ontologies and Data in Life Sciences, Koblenz, Germany, 2013, pp. 19–20.

[7] B. Smith *et al.*, “The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration”, *Nature biotechnology*, vol. 25, no. 11, pp. 1251–1255, 2007.

[8] International Organization for Standardization, “ISO/IEC 19510:2013 Information technology - Object Management Group Business Process Model and Notation.” Geneva:ISO, Nov-2013.

[9] H. Knublauch, J. A. Hendler, and K. Idehen, “SPIN-overview and motivation”, *W3C Member Submission*, 2011. [Online]. Available: http://www.w3.org/Submission/spin-overview/. [Accessed: 09-Jan-2017].

[10] ISO, “Information technology — Open Distributed Processing — Reference model: Overview”, Nov-2008. [Online]. Available: http://www.itu.int/rec/T-REC-X.901-199708-I/en. [Accessed: 09-Jan-2017].

[11] G. A. Uribe, B. Blobel, D. M. López, and S. Schulz, “A generic architecture for an adaptive, interoperable and intelligent type 2 diabetes mellitus care system”, *Stud Health Technol Inform*, vol. 211, pp. 121–131, 2015.

[12] G. A. Uribe, B. Blobel, D. M. López, and A. A. Ruiz, “Specializing architectures for the type 2 diabetes mellitus care use cases with a focus on process management”, *Stud Health Technol Inform*, vol. 211, pp. 132–142, 2015.

[13] G. A. Uribe, B. Blobel and D. M. López, “Systems architectures and ontologies applied to health information systems ”, 1st ed., Editorial Universidad del Cauca, Popayán, 2019.

[14] International Labour Organization, “International Standard Classification of Occupations (ISCO)”, *ISCO - International Standard Classification of Occupations*, 2007. [Online]. Available: http://www.ilo.org/public/english/bureau/stat/isco/. [Accessed: 06-Apr-2014].

[15] G. Uribe, “Type 2 diabetes mellitus Project”, 20-Sep-2018. [Online]. Available: https://github.com/gaurgo/DiabetesCare. [Accessed: 10-Sep-2019].

[16] J. S. Giboney, S. A. Brown, P. B. Lowry, et. al., “User acceptance of knowledge-based system recommendations: Explanations, arguments, and fit”, *Decision Support Systems*, vol. 72, pp. 1–10, 2015.