

A Systematic Approach to Define Requirements and Engineer the Ontology for Semantically Merging Data Sets for Personal-Centric Healthcare Systems

Aleksandr Kormiltsyn

Department of Software Systems, Tallinn University of Technology,
Akadeemia tee 15A, 12816, Tallinn, Estonia
alexandrkormiltsyn@gmail.com

Abstract. A fundamental paradigm shift to patient-centric health services is needed to meet the challenges such as increasing healthcare costs, ageing population, and unhealthy lifestyles. Designing patient-centric services is challenging because of medical- and health data heterogeneity and lack of standardization for Personal Health Records (PHR). Further, designing patient-centric systems, especially in healthcare is a challenge as they are influenced by emotions when users are influenced by emotions triggered by privacy issues. In this study, we focus on the engineering method to design a patient-centric system that considers medical-data heterogeneity and emotional security issues. We plan to develop a methodology for designing a patient-centric system with a multi-agent system (MAS) approach where agents operate with smart contracts. An emotional security framework is integrated into the design of such systems. Multi-faceted validation and verification are used to evaluate the engineered method. A framework for Evaluation in Design Science Research (FEDS) is used to define the evaluation activities and tools.

Keywords: eHealth, Patient-centric systems, Smart contracts, Emotional security, Multi-agent systems, Ontology

1 Introduction

Health information technology has the potential to improve medical quality, patient safety, educational resources and patient-physician communication while decreasing costs [7]. The development of central systems such as Smart Open Services for European Patients (epSOS)¹ and the Estonian Health Information System (EHIS)² increases interoperability on national- and international levels. Still, the available data is often fragmented and the secondary use is limited [11]. The term 'secondary use' can be defined as the use of data collected for one purpose to study a new problem. Secondary uses of data could include quality

¹ <http://www.epsos.eu>

² <http://www.e-tervis.ee>

measurement, public health surveillance, and patient access to data about their illness [6]. The fragmentation stems from decentralized and disintermediated data sets.

The main purpose of Personal Health Records (PHR) is that a patient is the author and owner of his/her medical data that can be shared with other individuals, including healthcare professionals, or automated clinical decision-support services. The growing amount of personal data available increases the role of patient-centric systems in healthcare. In contrast, Electronic Health Records (EHR) are created and owned by healthcare providers.

The complexity of data analysis, processing and security increases together with the number of different healthcare systems. Another problem is that the data is created by healthcare professionals and based on disease episodes that are only a part of the patients' health-related data and history. For example, several patients with chronic conditions measure their blood pressure, or blood glucose regularly while these data sets are not stored in the context of EHR and thus, limiting their future use.

Design and development of patient-centric systems are challenging because users are influenced by emotions triggered by privacy issues. As users are not forced to use such software, these issues lead to adoption failure, e.g., individuals do not trust to share their sensitive medical data with an application that does not provide transparency of personal data usage.

The importance for a healthcare domain is related to the possibility of sharing PHR between different stakeholders. Individuals could manage their own data and share it with healthcare providers and other stakeholders such as medical researchers, health insurance providers, and automated Clinical Decision Support Systems (CDSS). Additional medical data available to healthcare provider enables the provision of personalized guidelines to a patient. Personalized guidelines help patients to understand how to achieve their goals. On the other hand, the availability of CDSS reduces the number of people requiring direct contact with a doctor and increases the accessibility of medical services to people who really need it. Further, researchers gain access to personal medical data.

In this research, we focus on the PHR usage in the healthcare processes and involvement of patients as main responsible stakeholders by engineering a method for designing decentralized patient-centric systems. The main proposed method's concepts are: the number of healthcare processes is not limited; trust, transparency and influence of patient's emotions during medical data sharing are essential for the usage of such a system; medical data heterogeneity is unavoidable.

To enable these main concepts we use different techniques: Multi Agent Systems (MAS) to support an unlimited number of healthcare processes, smart contracts are used by both human and non-human agents to enable trust and transparency while sharing PHR, embedding emotional security framework in the MAS design phase to avoid negative emotions while using such a system. The ontology is used to describe medical data integration processes between different agents.

The remainder of the paper is structured as follows. Section 2 presents related work and Section 3 describes the problem statement as well as our contributions. Next, Section 4 presents the design-science research method research actions in details and Section 5 references the conference paper as initial results used as a basis for the current research. Section 6 provides the evaluation plan for designed methodology, Section 7 concludes the paper together with providing directions for future research. In a final Section 8 acknowledgements to supervisors are presented.

2 State of the Art

Literature review [3, 5, 22] shows that researchers provide solutions for integration challenges based on data-heterogeneity reduction by developing a new ontology or merging existing ones. This approach is effective for decreasing the heterogeneity of different EHR standards and a limited number of processes. However, the amount of PHR data is rapidly growing with the development of, e.g., IoT devices. The number of processes using PHR is not limited, as patients visit different hospitals. Therefore, an efficient integration of PHR and EHR requires a dynamic approach that focuses on the process rather than on data standardization. Paying attention to data-flow processes is an important part of EHR and PHR data integration as health data flows along processes in distributed systems. Research [2] proposes smart contracts for logging the patient-provider relationships that associate a medical record with viewing permissions and data retrieval instructions for the execution on external databases.

The role of emotions and associated quality goals that influence the adoption and effective use of socio-technical systems is detailed in research [15]. In order to design and develop a system, it is important to elicit requirements in the form of functional, nonfunctional and emotional goals [15]. To do this, some research has identified methodologies for segmenting the target users of a system as different stakeholders have different goals and requirements [21]. Research [21] hypothesizes that user segmentation is helpful in requirements engineering to direct systems analysts and owners to segment potential users and select target users as the first stage of requirements engineering. In [17], the authors introduce a straightforward notation for modeling emotional goals in agent-oriented modeling. Further, little is known about software engineering methodology defining how best emotional requirements can be elicited from stakeholders. There is no systematic methodology to help elicit such emotional requirements early on before the systems are designed.

3 Problem Statement and Contributions

Healthcare services become more expensive and their accessibility to patients decreases. Patient-centric healthcare systems enable PHR data collection and processing by CDSS and often do not require physical contact between patient and doctor. The PHR and EHR integration requires a solution for medical data

heterogeneity as PHR does not have common data standard. The processing of integrated PHR and EHR needs to be transparent for patients so they adopt the patient-centric system. In this paper, we fill two gaps: (1) the integration process of PHR and EHR; (2) and embedding an emotional security framework into the design of patient-centric systems. In the scope of integrated healthcare, there are different main stakeholders: individuals, researchers that use medical data, healthcare professionals, state organizational units as statistics and social departments, and insurance companies.

We develop a methodology for patient-centric systems design. The methodology includes an emotional attachment framework to engage users and alleviate the users from feeling a negative emotion such as distrust, lack of ownership private and confidential data. Lack of transparency in health data processing leads to negative feelings for an individual. Target users for this methodology are IT architects and designers because there is no current methodology in IT for designing patient-centric systems.

Our planned contribution is a methodology for the design of dynamic PHR and EHR integration processes. We plan to use intelligent agents in MAS and smart contracts for sharing PHR. Intelligent agents are used for PHR collection from personal devices and can provide feedback based on the collected data analysis while blockchains provide transparency of PHR usage to the patient. To eliminate emotional security risks, we expand traditional functional and non-functional requirements of software system with emotional ones.

The main research question is how to dynamically merge and process integrated PHR and EHR in the cross-organizational processes of the patient-centric, decentralized healthcare system? To answer this question, a number of challenges (sub-questions) need to be addressed.

The first sub-question is how to collect and process PHR data in the context of integrated healthcare to improve the quality of diagnostics and cures? To answer this question, we focus on the dynamic integration of EHR and PHR considered from a process-based point of view. After defining the dynamic integration process, we turn to the shared cross-organizational processes.

The second sub-question is how to specify cross-organizational processes for handling EHR and PHR in a dynamic and secure way in a distributed patient-centric systems? The dynamic integration of EHR with PHR and the embedded emotional security framework enables the effective adoption and usage of integrated data in a different number of e-services. Smart contracts used by intelligent agents in multi-agent systems (MAS) eliminate trust issues between stakeholders while providing a patient with PHR and EHR data security.

As the security of personal data is a very sensitive and important topic, the third research question is how to perform a privacy-aware analysis of integrated PHR and EHR in the context of the cross-organizational process in a distributed patient-centric system? The privacy aware protocol [19] that is used in data-driven EHR systems for private data analysis is extended to support integrated PHR and emotional security aspects. The protocol is based on the emotional

requirements to PHR data sharing as well as the ontology for the collection of emotional goals and secured views proposed in [19].

4 Research Methodology and Approach

For this research, the design-science research method is considered. This method is solution-oriented and requires the creation of an innovative, purposeful artifact for a special problem domain [24]. First, we consider the dynamic integration process for integrating the PHR with EHR. The proposed process is used as a foundation for designing cross-organizational processes where medical data is used by different human- and non-human agents. Finally, we define security protocol based on the emotional aspects of information security.

For designing dynamic integration processes, the requirements for PHR data collection are defined. The ontology for the PHR and EHR data integration and the integration process flow are presented. The requirements are defined with the Agent-Oriented Modelling (AOM) approach as in [23, 14]. According to [20], analyzing the problem of this socio-technical domain can be performed by using a goal model. The objective of goal models is to serve as communication media between technical- and non-technical stakeholders for generating understandable domain knowledge. We use goal models in the context of requirements-engineering processes.

The processing of integrated EHR and PHR requires an ontology definition to reduce semantic- and syntactic heterogeneity. An ontology for the dynamic integration process is designed using the Protégé tool³ together with the Web Ontology Language (OWL)⁴. OWL is used to describe taxonomies and classification networks. An evaluation of the ontology is performed with the Hermit⁵ reasoner that is based on a novel hypertext calculus [18] to provide much more efficient reasoning than any previously known algorithm.

We describe the PHR and EHR integration process with the Business Process Model and Notation (BPMN) [1], a graphical representation for specifying business process model. BPMN is suitable for reasoning about cross-organizational integration processes including PHR- and EHR integration process. The integration process is validated with the Signavio⁶ tool that checks BPMN diagrams for existing conflicts.

For the evaluation of defined integration process of PHR and EHR, we build a formal Colored Petri Nets (CPN) model in order to detect and eliminate eventual design flaws, missing specifications, security and privacy issues [8, 10]. A CPN is a graphical oriented language for the design, specification, simulation as well as the verification of systems and describes the states of a modeled system and the events (transitions) that cause the system to change states [13]. We refer the reader to [9] for more information about CPN.

³ <http://protege.stanford.edu>

⁴ <https://www.w3.org/2001/sw/wiki/OWL>

⁵ <http://www.hermit-reasoner.com/>

⁶ <https://www.signavio.com/>

To define the cross-organizational processes where medical data is used by different human- and non-human agents, first intelligent BDI agents are described as they offer a straightforward formalization of reasoning human agents with intuitive concepts (beliefs, desires, and intentions) that closely match human reasoning[4]. Next, best practices for designing decentralized healthcare systems are proposed.

We define best practices for designing decentralized healthcare systems by integrating smart contracts in the communication layer between different non-human agents involved in the cross-organization process. To describe non-human BDI agents, ontology built for the integration process in the previous step is extended using OWL and validate it with HermiT reasoner.

Emotional requirements for PHR handling and sharing processes are included in the early stages of the PHR and EHR data integration. The emotional framework [16] is integrated into the AOM approach and extends the AOM goal model with emotional goals in order to define the emotional requirements for PHR data sharing. The ontology for the integration process is extended to cover the collection of emotional aspects. We use BPMN to describe the security protocol and blockchain as a technology enabling transparency in the cross-organizational process.

To meet our goal of designing the architecture of a patient-centric system, three milestones are defined. First, the design of collecting and processing the PHR data, then the specification of a cross-organizational process followed by design of a secure PHR sharing protocol based on the emotional requirements. To define requirements for the PHR data collection, we use AOM. The ontology for the PHR and EHR integration process is defined using OWL and the integration process is described with BPMN. The specification of a cross-organizational process includes a definition of non-human agents described with AOM and OWL. The usage of smart contracts is considered in the context of the dynamic integration process. Secure PHR sharing protocol for the PHR includes a definition of emotional requirements for PHR data sharing described with AOM and the emotional framework, followed by the ontology for PHR emotional aspects collection described with OWL. Finally, we design a security protocol for PHR sharing using BPMN and blockchain technology.

5 Preliminary or Intermediate Results

In the workshop paper [12], we consider the dynamic integration of EHR and PHR from the process based point of view. The process workflow emphasizes the EHR and PHR context difference while the system requirements are defined with an AOM goal model describing functional and quality goals. The ontology for the integration process defined in the workshop paper is a foundation describing main concepts for the integration process. Finally, the BPMN representation of the integration process shows different stages of PHR- and EHR data preparation and -processing. All these findings set the scope of our research and we use them as a basis for future work.

6 Evaluation Plan

We use a combination of different approaches for the evaluation. First, CPN is used to evaluate processes in the design phase to eliminate, or minimize the security risks. Also, it is possible to consider concurrency conflicts, dependability issues and detect and eliminate eventual design flaws as well as security and privacy issues with CPN. For architecture design, we use a combination of Requirements Bazaar for requirements gathering and also for the evaluation of the architecture. The Architecture Tradeoff Analysis Method (ATAM) is used for evaluating and analyzing in a qualitative and empirical way the architecture of a patient-centric system with the related functional, non-functional and emotional requirements. ATAM is mostly applied for exploring the use of best-practice architectural styles, for exploring quality attributes of architectures and for evaluating existing systems. ATAM also helps to modify an architecture, or integration work with new systems. Also, we use the HerMiT reasoner for validating the ontology for the integration process and BPMN for reasoning about cross-organizational integration processes. We also conduct user-centered evaluations to capture user experiences with the design of the system to measure their level of engagement and emotional aspects. This evaluation we conduct using the emotional attachment framework as a practice lens to direct interviews and usability testing activities.

7 Conclusions

In this paper, we describe current issues and challenges in electronic healthcare, focusing on the need to create a systematic approach to designing a decentralized patient-centric system that meets these challenges. We then define three milestones to meet our goal of designing the architecture of such a system. To solve the PHR heterogeneity issue in healthcare processes, first, the design of collecting and processing PHR is proposed. However, the patient involvement in healthcare processes creates additional challenges for such processes and therefore, we present the specification for cross-organizational processes working with PHR. Shifting the paradigm to patient-centric increases the role of patient emotions and privacy issues. Thus, we design the secure PHR sharing protocol that enables transparency of the PHR usage to the patient and based on the patient's emotional requirements.

8 Acknowledgments

The author would like to thank both supervisors Alex Norta from Tallinn Technical University of Technology and Antonette Mendoza from the University of Melbourne.

References

1. Allweyer, T.: BPMN 2.0: introduction to the standard for business process modeling. BoD–Books on Demand (2016)
2. Azaria, A., Ekblaw, A., Vieira, T., Lippman, A.: Medrec: Using blockchain for medical data access and permission management. In: Open and Big Data (OBD), International Conference on. pp. 25–30. IEEE (2016)
3. del Carmen Legaz-García, M., Martínez-Costa, C., Menárguez-Tortosa, M., Fernández-Breis, J.T.: A semantic web based framework for the interoperability and exploitation of clinical models and ehr data. Knowledge-Based Systems 105, 175–189 (2016)
4. Casali, A., Godo, L., Sierra, C.: A graded bdi agent model to represent and reason about preferences. Artificial Intelligence 175(7-8), 1468–1478 (2011)
5. Dogac, A., Laleci, G.B., Kirbas, S., Kabak, Y., Sinir, S.S., Yildiz, A., Gurcan, Y.: Artemis: Deploying semantically enriched web services in the healthcare domain. Information Systems 31(4), 321–339 (2006)
6. Hersh, W.R.: Adding value to the electronic health record through secondary use of data for quality assurance, research, and surveillance. Clin Pharmacol Ther 81, 126–128 (2007)
7. Hoyt, R.E., Yoshihashi, A.K.: Health Informatics: Practical guide for healthcare and information technology professionals. Lulu. com (2014)
8. Jensen, K.: Coloured petri nets. In: Discrete Event Systems: A New Challenge for Intelligent Control Systems, IEE Colloquium on. pp. 5–1. IET (1993)
9. Jensen, K., Kristensen, L.M.: Coloured Petri nets: modelling and validation of concurrent systems. Springer Science & Business Media (2009)
10. Jensen, K., Kristensen, L.M., Wells, L.: Coloured petri nets and cpn tools for modelling and validation of concurrent systems. International Journal on Software Tools for Technology Transfer 9(3-4), 213 (2007)
11. Köpcke, F., Trinczek, B., Majeed, R.W., Schreiweis, B., Wenk, J., Leusch, T., Ganslandt, T., Ohmann, C., Bergh, B., Röhrig, R., et al.: Evaluation of data completeness in the electronic health record for the purpose of patient recruitment into clinical trials: a retrospective analysis of element presence. BMC medical informatics and decision making 13(1), 37 (2013)
12. Kormiltsyn, A., Norta, A.: Dynamically integrating electronic-with personal health records for ad-hoc healthcare quality improvements. In: International Conference on Digital Transformation and Global Society. pp. 385–399. Springer (2017)
13. Leiding, B., Norta, A.: Mapping requirements specifications into a formalized blockchain-enabled authentication protocol for secured personal identity assurance
14. Lister, K., Sterling, L., Taveter, K.: Reconciling ontological differences by assistant agents. In: Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems. pp. 943–945. ACM (2006)
15. Mendoza, A., Miller, T., Pedell, S., Sterling, L., et al.: The role of users emotions and associated quality goals on appropriation of systems: two case studies
16. Miller, T., Pedell, S., Lopez-Lorca, A.A., Mendoza, A., Sterling, L., Keirnan, A.: Emotion-led modelling for people-oriented requirements engineering: The case study of emergency systems. Journal of Systems and Software 105, 54–71 (2015)
17. Miller, T., Pedell, S., Mendoza, A., Keirnan, A., Sterling, L., Lopez-Lorca, A.A.: Emotionally-driven models for people-oriented requirements engineering: the case study of emergency systems. IEEE Transactions on Software Engineering (2014)

18. Motik, B., Shearer, R., Horrocks, I.: Hypertableau Reasoning for Description Logics. *Journal of Artificial Intelligence Research* 36, 165–228 (2009)
19. Nguyen, T.A., Le-Khac, N.A., Kechadi, M.T.: Privacy-aware data analysis middleware for data-driven ehr systems. In: *International Conference on Future Data and Security Engineering*. pp. 335–350. Springer (2017)
20. Norta, A., Mahunnah, M., Tenso, T., Taveter, K., Narendra, N.C.: An agent-oriented method for designing large socio-technical service-ecosystems. In: *2014 IEEE World Congress on Services*. pp. 242–249. IEEE (2014)
21. Sherkat, M., Miller, T., Mendoza, A.: Does it fit me better? user segmentation in requirements engineering. In: *Software Engineering Conference (APSEC), 2016 23rd Asia-Pacific*. pp. 65–72. IEEE (2016)
22. Sonsilphong, S., Arch-int, N., Arch-int, S., Pattarapongsin, C.: A semantic interoperability approach to health-care data: Resolving data-level conflicts. *Expert Systems* 33(6), 531–547 (2016)
23. Sterling, L., Taveter, K.: *The art of agent-oriented modeling*. MIT Press (2009)
24. Von Alan, R.H., March, S.T., Park, J., Ram, S.: Design science in information systems research. *MIS quarterly* 28(1), 75–105 (2004)