RESEARCH

A sample article title

Jane E. Doe[1*] and John R.S. Smith[1,2]

[*]Correspondence:
jane.e.doe@cambridge.co.uk
[1]Department of Science,
University of Cambridge, London,
UK
Full list of author information is
available at the end of the article

Abstract

First part title: Text for this section.

Second part title: Text for this section.

Keywords: sample; article; author

Introduction

Data and Information modeling in the healthcare domain have witnessed significant improvements in the last decade owing to advances in the development of state-of-the-art information and communication technologies (ICT) and formalization of storage and messaging standards. Subsequently, the scope of Healthcare Management Information Systems (HMIS), medical ontologies, and Clinical Decision Support Systems (CDSS) has broadened, beyond the operational capabilities of traditional rule based systems. One of the major reasons behind this limitation is due to the numerous heterogeneities in healthcare at data, knowledge, and process level. Thus, healthcare interoperability which aims to provide a solution to this problem, can be compartmentalized into data interoperability, process interoperability, and knowledge interoperability. Data interoperability resolves the heterogeneity between data artificats, to enable, seamless and interpretable communcation among source and target organizations, while preserving the data's original intention during storage, communication, and usage (as defined by IEEE 610.12 [1], Health Level Seven International HL7, and Healthcare Information and Management Systems Society HIMSS [2]). On the other hand, process interoperability regulates the communication among organizational processes to provide compatability between process artifacts within and seamless transformations across different organizations[3]. Lastly, knowledge interoperability provides a sharing mechanism for reusing interpretable medical knowledge, acquired through expert intervention and other mechanisms, across decision support systems [4]. In more tangible terms, healthcare interoperability at data, process, and knowledge level can be exemplified within the healthcare constraints experienced due to the emergence of Covid 19. The operational capabilities of the current healthcare service delivery infrastructure has gone under tremendous stress due to Covid 19. World over, large primary healthcare units have managed to create separate units for managing patients, suffering from extreme cases of the novel coronavirus. For secondary and tertiary care units, government involvement has become necessary to filter coronavirus patients and adhering to a national pandemic response policy. These complex circumstances have enhanced the need for sharing patient data and state-of-the-art medical knowledge in real-time, to provide the medical experts with a tool to make accurate and timely decisions. Data interoperability can enable the front line medical workers to fetch, understand, and use patient data, especially comorbidities, across organizational and physical boundaries, without suffering from societal taboos that may prevent the patient from sharing their complete and accurate medical histories. Knowledge interoperability can improve the knowledge acquisition and sharing protocols to provide the medical experts such as epidieomologists and vaccinologist, with latest information on affected population trends, disease diagnosis, treatment, and followup procedures, and interpretable decisions leading to positive or negative outcomes. Process interoperability can help reduce and in some cases remove the operational redundancies between health centers. In this way, successive healthcare treatments can take benefit from earlier diagnosis, treatment, and followup procedures, thereby reducing the stress on healthcare experts and systems. Standards such as Health Level Seven (HL7) Fast Healthcare Interoperability Resources (FHIR), and openEHR provide the foundations for storing and communicating medical data, through the use of well defined protocols. While systematized nomenclature of medicine—clinical terms (Snomed-CT) [5] and logical observation identifiers names and codes (LOINC) [6] provide a standard definition for clinical terminologies and laboratory tests, respectively. Similarly Medical Logic Module (MLM) provides a standardized way for expressing medical knowledge. However, the plethora of standards, necessitates the creation of bridging standards, that can resolve the heterogeniety between the medical standards. Substantial effort has gone into this endeavor with the Clinical Information Modeling Initiative (CIMI) [7] taking the lead in bridging the gap between HL7v3 and openEHR. Similarly, SNOMED CT and LOINC are working to resolve the redundancies between the two terminological standards since 2013. This healthcare interoperability solution follows a formal, albiet long process, which is greatly dependent on the human factor. However, the current healthcare scenario, requires a quick solution to create a scaffolding of an interoperable bridge between various healthcare providers. It is also important to ensure that this scaffolding should be able to support the formal standardization processes of the future. In [8]we have presented the Ubiquitous Health Platform (UHP), which provides semantic reconciliation-on-read based data curation for resolving data interoperability between various schema. This methodology is based on the creation and management of schema maps, that can provide the framework for transforming a source schema into a target schema. In the current manuscript, we will present our research work to build and manage the schema map knowledge base. Overall, our methodology has two major portions, firstly we apply a novel schema matching technique to create a transformation function σ for the participating legacy schema, and secondly, we have used the Ripple Down Rules (RDR) to manage our knwoledgebase, which will be presented in some detail, focusing primarily on the search and evolution services. In particular,

– Section 2 contains the details of our methodology, where we aim to present a reproducible theoretical framework.
– Section 3 provides the experimental setup
– Section 4 presents the results
– Section presents the related work
– Section 6 concludes the paper.

Related Work

Althubait et al. [9] proposed an ontology expansion methodology that identifies and extracts new class from text articles using word embedding and machine learning techniques. The authors identified the similarity of tokens and phrases of the text articles with the exiting classes of the ontology. The target ontology is expanded with classes from text articles having greater similarity with that of already added classes. A similar word embedding technique was also used by Nozaki et al. [10], where the authors used instance based schema matching technique to identify the semantic similarity between two instances. The results of the study showed the possibility of detecting similar string attributes of different schemas. Yousfi et al. [11] also utilized semantic base techniques and proposed xMatcher XML schemas matching approach. xMatcher transforms schemas into a set of words, followed by measuring words context, and relatedness score using WordNet. The terms from different schemas having similarities greater or equal to 0.8 are considered similar. Bylygin et al. [12] devised an ontology and schema matching approach by combining lexical and semantic similarity with machine learning approaches. The authors used lexical and semantic measures as features and trained various machine learning algorithms including Naive Bayes, logistic regression, and gradient boosted tree. The result achieved showed that the combination of algorithms outperformed the single modal.

Martono et al. [13] provided overview of previous studies related to linguistic approaches used for schema matching. Linguistic methods focused on finding strings and evaluate there similarity in different schemas. The string are normally normalized before to align both the strings before similarity comparison. The normalized strings are categories based on the information relatedness and element with similar category are compared using various similarity measure including Jaro-distance, Lavenstein (edit-distance), and many more. Alwan et al. [14] have summarized the techniques used in the literature for schemas and instances based schema matching. The information used for schema matching is categories into schema information, instance and auxiliary information. Most of the searchers have used syntactic techniques (including n-gram, and regular expression), semantic techniques (including Latent Semantic Analysis, WordNet/Thesaurus and Google Similarity) for schema level and instance level matching to achieve the final goal of data/information interoperability. Kersloot et al. [15] performed a comprehensive systematic review to evaluate natural language processing (NLP) algorithms used for clinical text mapping onto ontological concepts. The findings of the studies were evaluated with respect to five categories; use of NLP algorithms, data used, validation and evaluation performed, result presentation, and generalization of results. The authors revealed that over one-fourth of the NLP algorithms used were not evaluated and have no validation. The systems that claimed generalization, were self evaluated and having no external validation.

Acknowledgements
Text for this section...

Funding
Text for this section...

Abbreviations
Text for this section...

Availability of data and materials
Text for this section...

Ethics approval and consent to participate
Text for this section...

Competing interests
The authors declare that they have no competing interests.

Consent for publication
Text for this section...

Authors' contributions
Text for this section ...

Authors' information
Text for this section...

Author details
[1]Department of Science, University of Cambridge, London, UK. [2]Institute of Biology, National University of Sciences, Kiel, Germany.

References
1. Geraci, A., Katki, F., McMonegal, L., Meyer, B., Lane, J., Wilson, P., Radatz, J., Yee, M., Porteous, H., Springsteel, F.: IEEE Standard Computer Dictionary: Compilation of IEEE Standard Computer Glossaries. IEEE Press, ??? (1991)
2. Healthcare Information and Management Systems Society: Definition of Interoperability (2013). http://www.himss.org/library/interoperability-standards/what-is
3. Khan, W.A., Hussain, M., Latif, K., Afzal, M., Ahmad, F., Lee, S.: Process interoperability in healthcare systems with dynamic semantic web services. Computing **95**(9), 837–862 (2013)
4. Ali, T., Hussain, M., Khan, W.A., Afzal, M., Hussain, J., Ali, R., Hassan, W., Jamshed, A., Kang, B.H., Lee, S.: Multi-model-based interactive authoring environment for creating shareable medical knowledge. Computer Methods and Programs in Biomedicine **150**, 41–72 (2017)
5. SNOMED Clinical Terminologies. http://www.snomed.org/snomed-ct/five-step-briefing Accessed 2020-03-19
6. LOINC. https://loinc.org/ Accessed 2019-03-19
7. CIMI: Clinical Information Modeling Initiative (CIMI) (2015). http://www.opencimi.org/ Accessed 2020-03-19
8. Satti, F.A., Ali, T., Hussain, J., Khan, W.A., Khattak, A.M., Lee, S.: Ubiquitous Health Profile (UHPr): a big data curation platform for supporting health data interoperability. Computing (2020). doi:10.1007/s00607-020-00837-2
9. Althubaiti, S., Kafkas, Ş., Abdelhakim, M., Hoehndorf, R.: Combining lexical and context features for automatic ontology extension. Journal of biomedical semantics **11**(1), 1–13 (2020)
10. Nozaki, K., Hochin, T., Nomiya, H.: Semantic schema matching for string attribute with word vectors. In: 2019 6th International Conference on Computational Science/Intelligence and Applied Informatics (CSII), pp. 25–30 (2019). IEEE
11. Yousfi, A., El Yazidi, M.H., Zellou, A.: xmatcher: Matching extensible markup language schemas using semantic-based techniques. International Journal of Advanced Computer Science and Applications **11**(8), 655–665 (2020)
12. Bulygin, L.: Combining lexical and semantic similarity measures with machine learning approach for ontology and schema matching problem. In: Proceedings of the XX International Conference "Data Analytics and Management in Data Intensive Domains"(DAMDID/RCDL'2018), pp. 245–249 (2018)
13. Martono, G.H., Azhari, S.: Review implementation of linguistic approach in schema matching. International Journal of Advances in Intelligent Informatics **3**(1), 1–9 (2017)
14. Alwan, A.A., Nordin, A., Alzeber, M., Abualkishik, A.Z.: A survey of schema matching research using database schemas and instances. International Journal of Advanced Computer Science and Applications **8**(10) (2017)
15. Kersloot, M.G., van Putten, F.J., Abu-Hanna, A., Cornet, R., Arts, D.L.: Natural language processing algorithms for mapping clinical text fragments onto ontology concepts: a systematic review and recommendations for future studies. Journal of biomedical semantics **11**(1), 1–21 (2020)

Figures

Figure 1 Sample figure title

Figure 2 Sample figure title

Table 1 Sample table title. This is where the description of the table should go

	B1	B2	B3
A1	0.1	0.2	0.3
A2	...	..	.
A3	..	.	.

Tables

Additional Files

Additional file 1 — Sample additional file title

Additional file descriptions text (including details of how to view the file, if it is in a non-standard format or the file extension). This might refer to a multi-page table or a figure.

Additional file 2 — Sample additional file title

Additional file descriptions text.