Intelligent Automation And Soft Computing, 2020
Copyright © 2020, TSI® Press
Vol. 26, no.2, 237–251
https://doi.org/10.31209/2019.100000151

A Fuzzy Ontological Infrastructure for Semantic Interoperability in Distributed Electronic Health Record

Ebtsam Adel[1], Shaker El-sappagh[2], Mohammed Elmogy[3], Sherif Barakat[1], Kyung-Sup Kwak[4]

[1]Information Systems Dept., Faculty of Computers and Information, Mansoura University, Egypt
[2]Information Systems Dept., Faculty of Computers and Informatics, Benha University, Egypt
[3]Information Technology Dept., Faculty of Computers and Information, Mansoura University, Mansoura 35516, Egypt
[4]Department of Information and Communication Engineering, Inha University, South Korea.

ABSTRACT
Information technology is a beneficial tool for the healthcare industry. Health informatics is concerned with using ICT within the healthcare system. Different electronic health record (EHR) systems independently store large amounts of medical data in various structures and formats. Achieving semantic interoperability in EHR environments will improve the healthcare industry. In our previous studies, we proposed a framework that identifies the different heterogeneous medical data sources. In this paper, we move towards implementing the first module of that framework. We expect our framework to be a step towards improving performance and reducing both human mediation and data losses.

KEYWORDS: EHR semantic interoperability, XML, LinkEHR, XML2OWL, fuzzy ontology.

1 INTRODUCTION

SEMANTIC interoperability (SI) is one of the most critical and serious problems in distributed electronic health records (EHRs). An EHR is a combination of patient's linked data which permits efficacious sharing of that related data. EHR interoperability is important in improving healthcare quality. IT helps in delivering the correct information at a suitable time. Moreover, it reduces the cost of healthcare and supports correct decisions. Interoperability is defined by IEEE as "the ability of more than one component or system to interchange data and information and to utilize that exchanged information" (IEEE Standard Glossary of Software Engineering Terminology, 1990). There are three main layers of interoperability (Kubicek et al., 2011), as illustrated in Figure 1.

SI is an important requirement between various systems, such as hospital information systems and mobile applications. IN the absence of semantic interoperability; many different medical data will be in

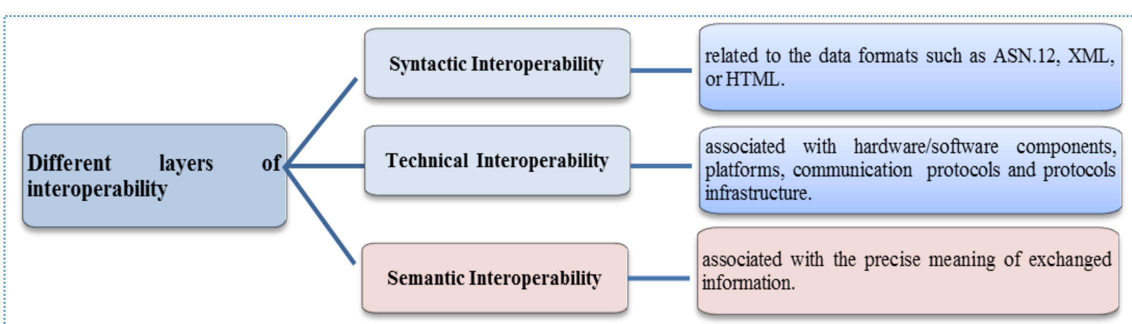


Figure 1. The main EHR interoperability layers.

CONTACT Kyung Sup Kwak kskwak@inha.ac.kr

isolation. Therefore, massive amounts of analytics and decision-making data would be lost. SI helps in ensuring that the received system can understand and interpret the sent information without ambiguity (National Health Service - NHS England, 2015). Achieving interoperability in an EHR environment presents many difficulties and challenges (The Health Information Technology Policy Committee, 2015). Heterogeneity is the biggest problem among those challenges. There are many kinds of heterogeneity due to differences in technologies, such as system software, hardware, and communications systems. Heterogeneous environments of healthcare systems ought to integrate with the different specifications of terminologies, data models, and architecture of various standards (ASTM E1238/1384, HL7 RIM, ISO/TS 18308:2004, OpenEHR, etc.). Semantic heterogeneity happens when there is a difference in meaning, translation, or proposed utilization of the same or related data (Sheth and Larson, 1990). In this paper, we move towards achieving the full semantic interoperability in distributed EHRs based on semantic web technologies using ontologies.

This paper is structured as follows. Various technologies and definitions discussed and manipulated in the rest of the paper will be defined in Section 2. Section 3 contains some studies and searches that deal with the integrating problem of heterogeneous medical data sources in distributed EHRs. In Section 4, a unified fuzzy ontology framework for a distributed EHRs semantic interoperability is recommended. This paper will be concluded in Section 5.

2 DEFINITIONS

THIS section contains some of the basic definitions for terms that are discussed later in this paper. It begins with the key Extensible Markup Language (XML) technology. Then it gives the reader a short introduction to the database. It contains a brief background about EHR standards and Archetype Definition Language (ADL) archetypes. It also provides details about semantic web technologies and ontologies.

2.1 XML

XML is the most powerful and familiar semi-structured data source. An XML document is a tree structure linearization. Several character strings are there at every tree's node. The character strings and the tree structure compose the XML contents (W3C, n.d.). According to W3C, there are two main XML schema formats: DTD and XSD. XML has a wide agreed-upon community for exchanging and storing data amongst different standards, providing many advantages (Thi et al., 2009). From those advantages are (1) flexibility of usage and simplicity (2) a common syntax for different systems, (3) a human-readable and easy-to-understand language, (4) versatility (as a key advantage) (Brewton et al., 2012), and (5) compatibility with many different object-oriented programming languages such as Java, Python, and C++. However, it supports neither semantics nor reasoning (Yahia et al., 2012). Ontology has the ability to support both reasoning and semantics with domain knowledge effectively and powerfully. Therefore that problem could be solved by translating XML documents into OWL or RDF.

2.2 Relational Database

A database is a related data collection. DBMS is the software able to manage and control it. The database consists of three levels of abstraction: conceptual, physical, and external designs (Ramakrishnan and Gehrke, 2009). The conceptual schema is used to depict the stored data. The physical schema describes how the relations are stored. External schemas allow data access at the user level. A relational database provides the most efficient storing and retrieving data technique, plus scalability and easy backup. An idiomatic definition of ER model could get in (Calvanese et al., 1999).

Despite all the mentioned capabilities of the database, nowadays, there is an urgent need to represent information and knowledge semantically to be machine readable. At the same time, the database suffers from weak semantics. As most data are stored and represented in a database, there are more trials to relational databases to semantic web technologies for getting more expressive information. Ontologies have many advantages over databases or any other source because they use more excellent language for expressing the information itself with natural expressions. The scientific community generally believes ontologies are the best technique for outlining and representing the truth because of their capacity for semantic concepts modeling (Martinez-Cruz et al., 2012). That is because they are represented with logic languages, as description logics (DL) or first-order logic. Also, ontologies use reasoners that can generate new information far from whether the data is defined therein or not (Martinez-Cruz et al., 2012).

2.3 EHR Standards, Archetypes, and Archetype Definition Language

ISO identified standard as "a document drawn up by consensus and confirmed by a recognized body. Its main aim is to achieve the best class in the order in a given context. It provides common guidelines, rules, and characteristics for activities or their results" ("ISO reference definitions - guide 2 - 2004 - rev," n.d.). Standards play a critical role in facilitating and achieving semantic interoperability in EHR environments. That is necessary to progress the care quality and to support safety. At the point when interoperability becomes ordinary, patients, clinicians, and scientists will appreciate secure access to the correct data at the ideal time and in the perfect place;

they will settle on more soundly based choices, prompting better patient results and fewer missteps. Interoperability is essential for procedural reengineering that will lessen pointless costs, costly mistakes, delays, and worthless reiteration (Benson, 2016). Many organizational standards designed to achieve interoperability in the EHR environment, including HL7, CIMI, CEN/ISO 13606, FHIR, CDISC, DICOM, IHE, and openEHR as depicted in Figure 2. Some challenges and constraints when using standards will be found in (Adel et al., 2018).

An archetype is a reusable and computable constraint statement of a specific domain relied on a reference model (a brief description is depicted in Figure 3). OpenEHR archetypes rely on the reference model of openEHR (Beale and Heard, 2007). They can represent EHR knowledge in the healthcare domain by referring to clinical concepts as biochemistry results, physical examination, laboratory test, and blood pressure. The main advantage of archetypes is that they are reusable and shareable, providing an interoperable way to manage the creation of data, effectiveness, and querying by ensuring that data comply with particular semantic and structures constraints (Martínez-Costa et al., 2009). Most recommendations consider that model as a convenient solution towards achieving EHR semantic interoperability by normalizing the transferred information between heterogeneous healthcare systems (Tapuria et al., 2013) (Maldonado et al., 2012). That is due to their ability to take into consideration the freedom of the reference model and to deal with different EHR architectures.

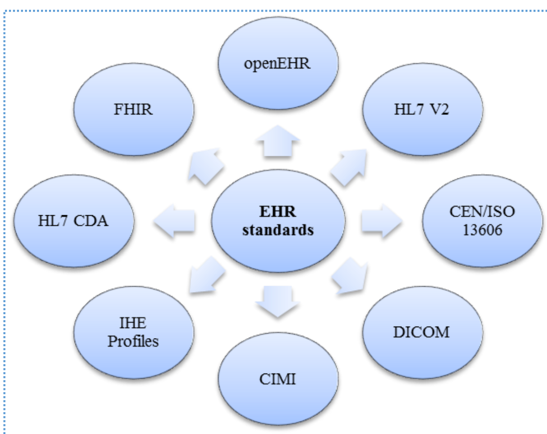


Figure 2. **The main EHR standards.**

Archetypes are considered a meeting point amongst EHR data sources and semantic-driven modeling (Maldonado et al., 2012). Archetypes are built in a hierarchical structure to formalize the natural tree of EHR data. Archetype nodes are distinguished by identifiers of semantic, that serve as the base for human-readable *meanings (Beale and Heard, 2007)*.

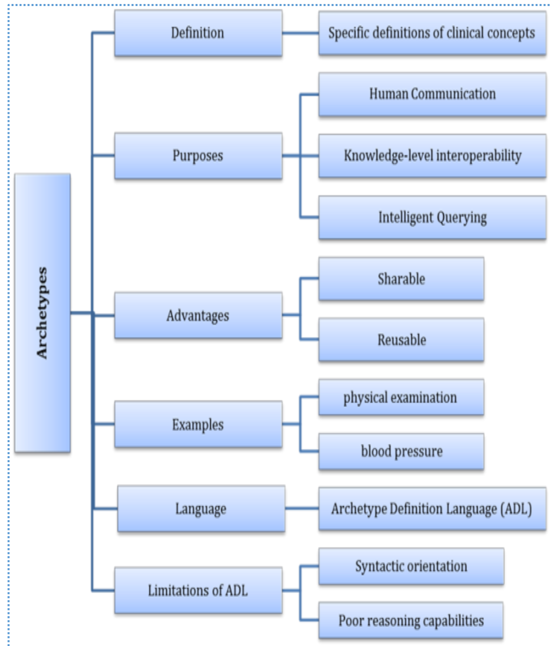


Figure 3. **A brief description of the archetype.**

2.4 The Semantic Web, the Ontology, and the Fuzzy Ontology

Berners et al. (2001) defined the semantic web as follows: " is a new current-web expansion in which the user can obtain a clear meaning of information and also enable people and computer-devices to work jointly together." The main purpose of the semantic web is to get the precise meaning for the exchanged information, and therefore, to make the shared documents human- and machine-readable at the same time. Semantic web technologies seem to be appropriate in order to achieve interoperability in EHR environments (Martínez-Costa et al., 2009). That is because semantic web technologies facilitate information integration of very different systems, just as they offer enough expressiveness and have the capability of automated reasoning. In addition, operations as classification, selection, comparison, and checking of consistency could be completed over OWL in a smoother and more effective method than ADL. RDF, XML, OWL, and OWL 2 are the main basic web technologies. The ontology is considered the cornerstone item of semantic web technologies; a brief description of the ontology is shown in Figure 4.

Mathematically, ontology model O can be defined as follows (Dadjoo and Kheirkhah, 2015):

Definition 1. O =(Cls,OP,SP,DP,HR,IFP,FP,I) where:

- Cls is a group of classes or concepts in the ontology. OP is an object property between two different concepts.
- SP is a sub-property of an object property.
- HR is a relationship hierarchy between two concepts.

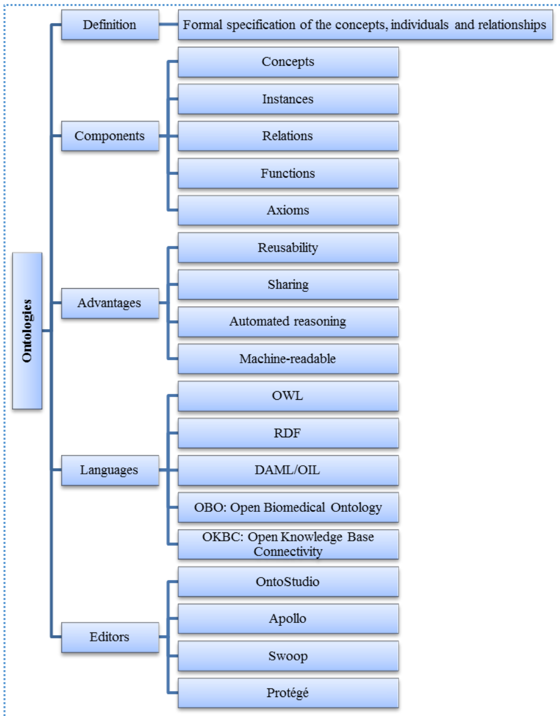


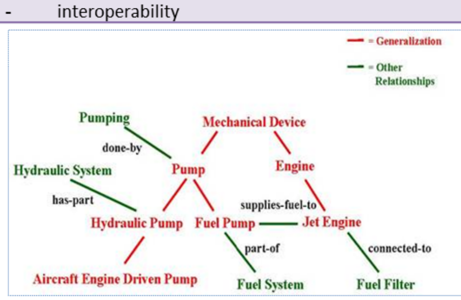
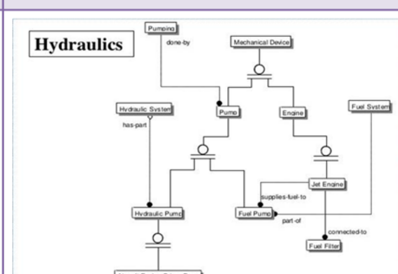
Figure 4. A brief description of the ontology.

- DP is a data type of property between literal value and a concept.
- FP is a functional property between literal value and a concept.
- IFP is an inverse functional property.
- I represents a group of ontology's instances.

Ontology is different from a database. Table 1 presents a comparison between ontology and database. The initial step in constructing any successful system for knowledge representation is to build a powerful ontological analysis of the domain or the field (Chandrasekaran and Josephson, n.d). That is because of the following capabilities.

1. Ontologies are enabled by terminology to work for coherent reasoning purposes.
2. Ontologies have the capability to share knowledge.
3. Shared ontologies enable building particular knowledge bases that depict specific situations, such as collecting data from many different categories.

Table.1 A comparison between ontology and database.

Criteria	Ontology	Database
Definition	"The shared understanding of a specific domain that is concurred between various representations. Such understanding encourages precise and viable communications of meanings, which leads to different benefits like sharing interoperability and reusing" [22].	"Is an organized data collection that could be searched systematically to maintain and retrieve information"
Focus	Focus on meaning	Focus on data
Similarities	- Use of relationships. - Relationship cardinality and domain constraints - Inheritance – define sub-type relationships	
Differences	- Semantic expressivity - Query flexibility - Interoperability - Accuracy of representation	- Security - Performance of handling data
When I need	- Semantically heterogeneous sources - Acquire new information - interoperability	- handle big amount of data - query time is required
Example		
Editors examples	Protégé, Apollo, Swoop, OntoStudio, and TopBraid Composer.	PostgreSQL, MySQL, Microsoft Access, SQL Server, Oracle, FileMaker, dBASE, and FoxPro.

Formal ontologies have the ability to support the automatic processing and recognition of heterogeneous expressions (Schulz and Martínez-Costa, 2013) (Mahalingam and N.Huhns 2000). An ontology involves two main components: taxonomy (classes and relationships) and inference rules (Kuck, 2004). There are several languages for building it. The most popular one is the Web Ontology Language, which depends on DL. There are many ontology editors, such as OntoStudio, Apollo, Swoop, Protégé, and TopBraid Composer. OntoStudio and Protégé are the most graphical ontology editors. Protégé is also the most widely used ontology editor and is more scalable and extensible. In this paper, we move towards introducing a unified framework based on fuzzy-ontology for semantic interoperability in distributed EHRs. A fuzzy ontology is a quintuple, F =< I, C, T, N, X > (Tho et al., 2006), where

- I is a group of objects (instances of the concepts).
- C is a group of fuzzy concepts.
- T refers to the fuzzy taxonomy relations between concepts.
- N refers to a group of non-taxonomy fuzzy relationships which has a tree structure for the instances.
- X is the axioms set.

The ontologies based on fuzzy-theory have many capabilities that exceed the traditional one (Díaz et al., 2014) such as:

1. It allows us to reason about and model incomplete, uncertain, and vague knowledge.
2. It can communicate with different applications interfaces.
3. It has the ability to interact with both unstructured and numeric data to ease the facts and rules expression.

3 RELATED WORK

THROUGH the last few years, there have been some trials that deal with the integrating problem heterogeneous medical data sources in distributed EHRs into a unified and interoperable model. From those trials, ontologies were selected as a common, usable, and shareable model. From the works based on semantic web technologies and ontologies, El Hajjamy et al. (2018) developed a semi-automatic ontology integration model. As shown in Figure 5, that model converted the classical data sources (RDB, XML, and UML) to local ontologies (OWL2), and then combined those ontologies into a global ontological one relied on structural, semantic and syntactic similarity measurement algorithms. Those techniques avoid redundancy in the output ontology.

Berges et al. (2012) presented an ontology-based framework that achieved interoperability between medical diagnosis statements from health information systems (Figure 6). The proposed architecture had the following features: it interacted with modules that aim to get clear representations for EHR information based on ontology and is considered the path mapping which is necessary for mapping axioms between ontological terms.

Kiourtis et al. (2017) proposed a multi-step nonexclusive semantic design that can be executed with different medical standards for productively managing heterogeneous EHRs' information. The proposed design consolidated a mechanism for extracting domain-specific information from classified EHR datasets, changing them into a common health language (CHL) through ontologies, while unknown datasets are converted and mapped into CHL utilizing ontology mapping techniques.

Myłka et al. (Myłka et al., 2012) presented an ontological system called X2R for consolidating heterogeneous data sources. Its main goal was to create a unified view of information stored in XML, relational, and Lightweight Directory Access Protocol (LDAP) data sources described in RDF format utilizing a common ontology according to a group of integrity constraints. The X2R system can work with a large dataset for a prolonged amount of time.

González et al. (2011) proposed an ontology-based framework for exchanging information and knowledge with a personal health record system and an EHR laboratory system. The architectural framework relies on GCM (*Generic Component Model*) (Blobel, 2002), as shown in Figure 7. That framework composed of three main systems: OpenMRS, Indivo, and the Bika laboratory information management system (BikaLIMS).

4 THE PROPOSED FRAMEWORK

IN any distributed and heterogeneous EHRs; the main means for achieving semantic interoperability is to establish a single and unified data model. In this paper, we recommend a fuzzy ontology united framework for handling data interoperability and integration problem between heterogeneous systems. The framework consists of three modules: local ontologies construction, global ontology construction, and the user application interface.

As shown in Figure 8, the proposed framework is built on the idea of converting each input source into an ontology representation. The main benefit of converting is to permit the independent development of the ontology source. Therefore, the increment and decrement of other input sources can be easily accomplished, and the integration process can be easily achieved. A crisp ontology can solve that problem to a certain degree. We look forward to extending the crisp ontology to a fuzzy one. We expect by doing that; the proposed framework will produce a more responsible, precise, and global semantic interoperable EHR environment. The architecture of framework composed of three base layers. The lowest one stores the EHR heterogeneous data formats. Those sources might be different

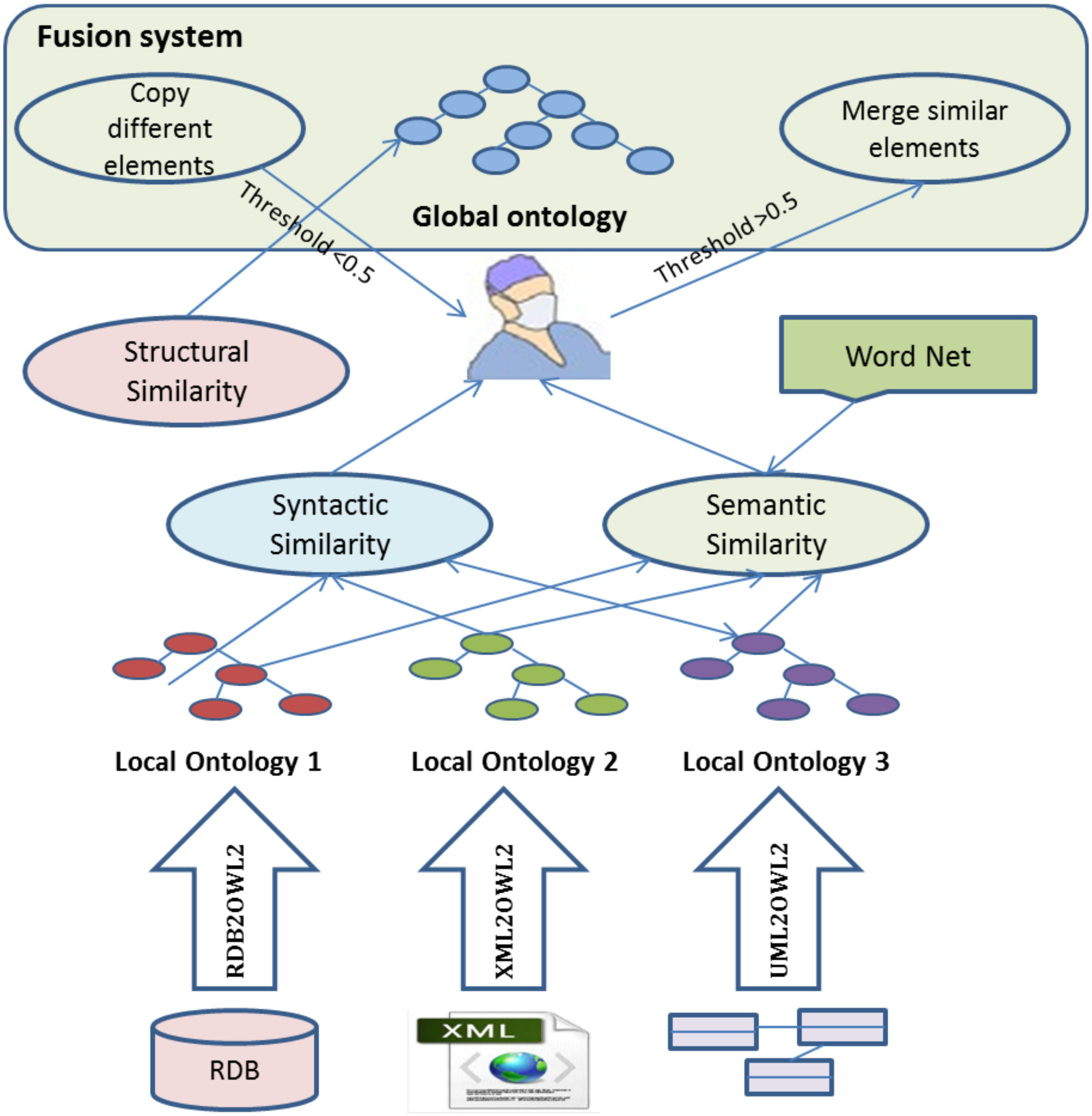


Figure 5. The proposed Hajjamy approach (El Hajjamy et al., 2018)

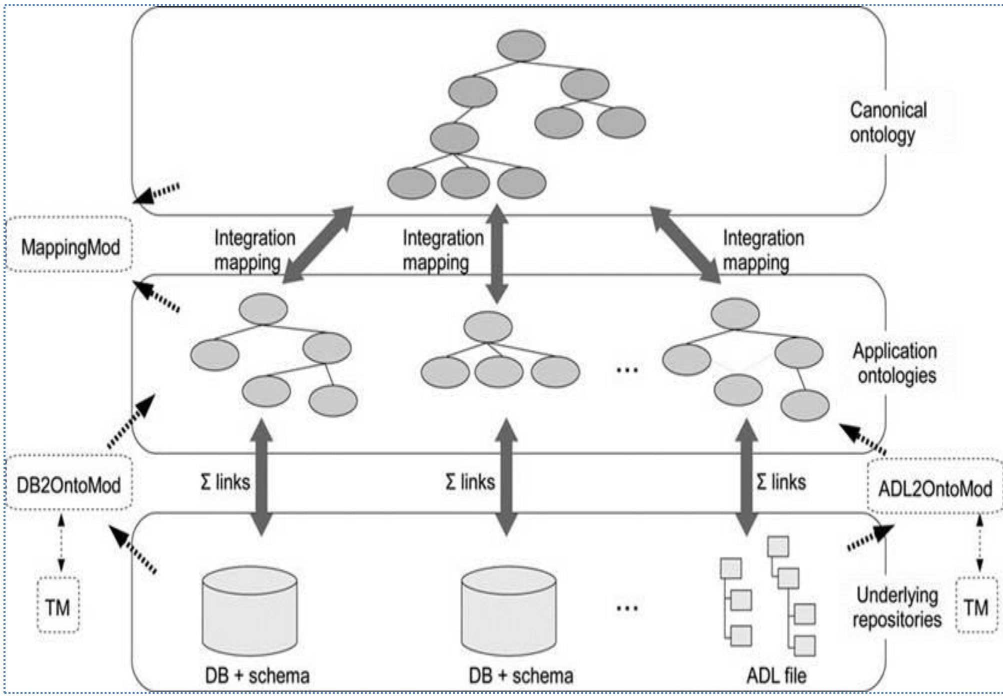


Figure 6. **Berges et al. (2012) global architecture.**

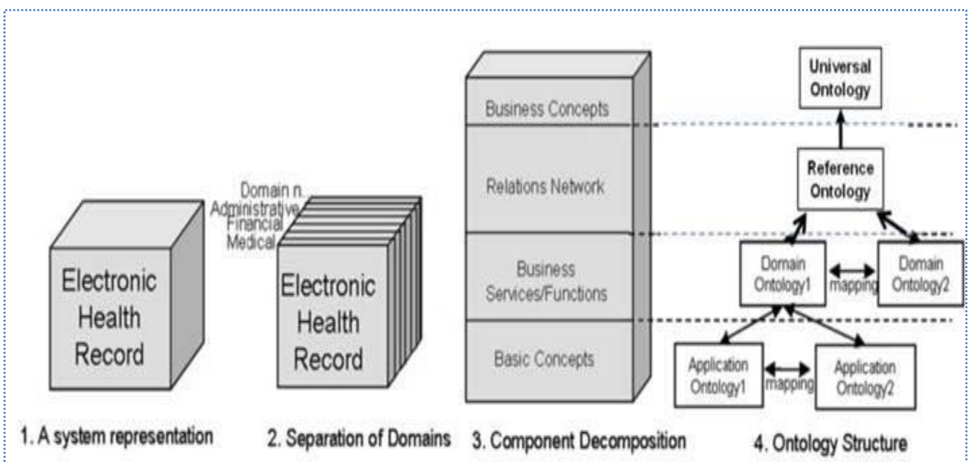


Figure 7. **GCM architectural design.**

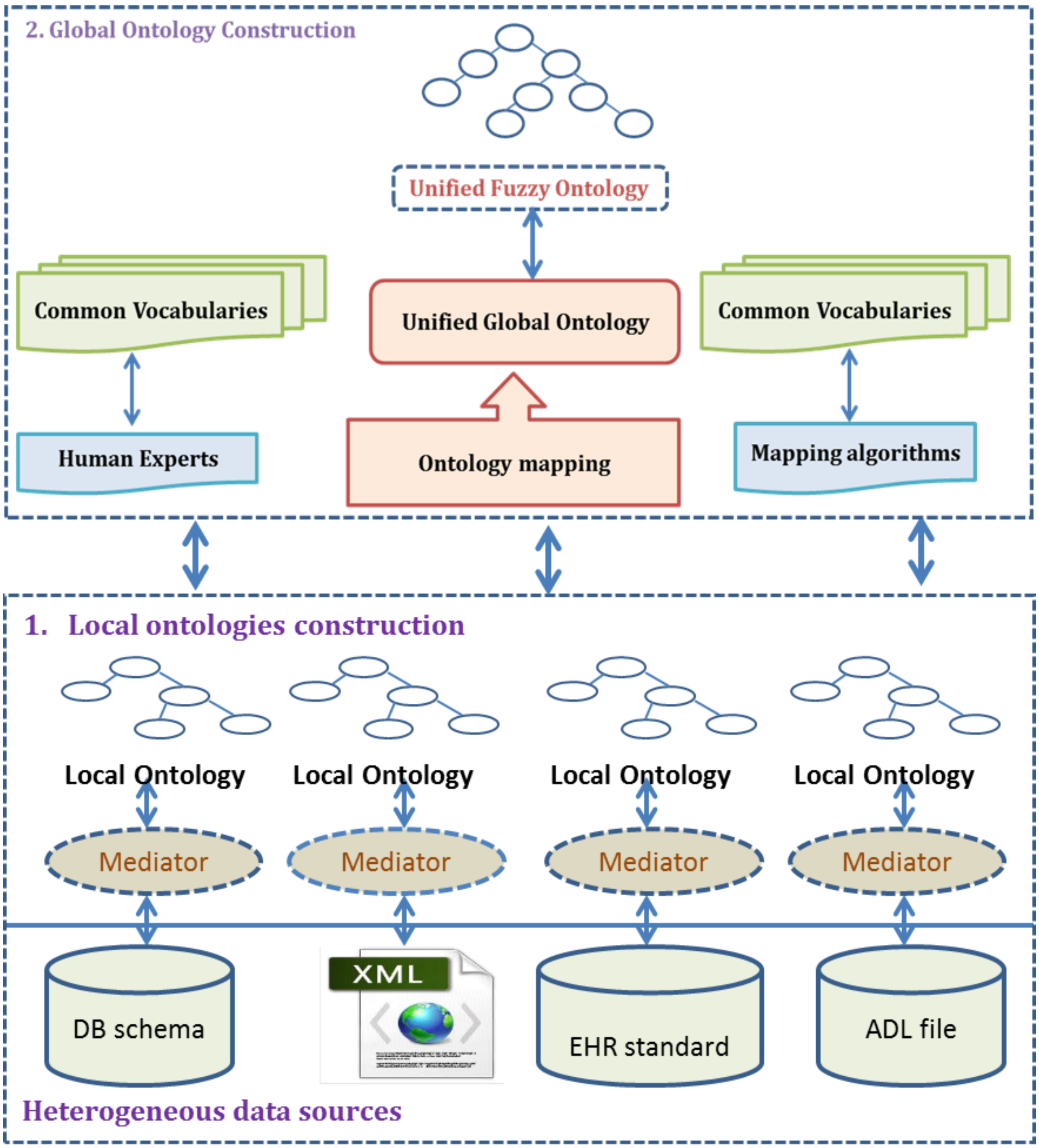


Figure 8. The proposed fuzzy ontology framework for distributed electronic health records.

databases with diversified schemas, XML files, EHR standards, CSV spreadsheet files, or ADL files.

The various inputs are converted into a local ontology using a suitable mapping tool for each type (that local ontology describes the local data), as depicted in Figure 9. In the middle layer; the locally constructed ontologies are mapped to a crisp global ontology. The global ontology constructed of all local ontologies and subsequently describes all data. After that, the crisp ontology is transformed into an integrated fuzzy ontology. The resulting ontology is capable of answering physicians' semantic queries at the interface layer. We foresee by using ontology and fuzzy ontology capabilities; the problem would be solved.

4.1 Ontologies Generation

Semantic interoperability is becoming one of the most important information technology topics. In the healthcare environment, heterogeneous systems ought to communicate with each other to provide complete

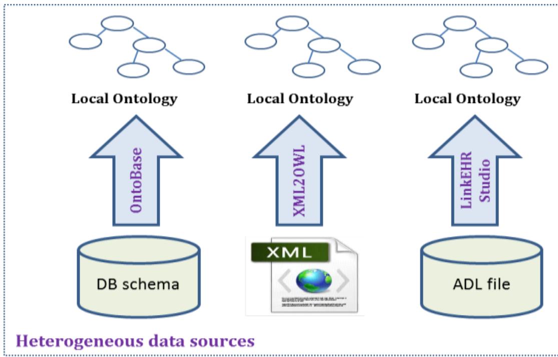


Figure 9. Local ontologies construction layer.

data at a suitable time. A specialist makes a query regarding all these different syntactically and semantically distributed data. The most complicated problem here is that; these sources use different vocabularies and synonyms, which may lead to semantic differences. The ontology-based data integration is not a new topic in the data integration research field. In this paper, we introduce a fuzzy ontology model to integrate all heterogeneous data sources. We concern on the first module and its software platforms of that proposed framework. In that module, we try to develop local ontologies from each input data source, like databases, XML files, and EHR standards.

4.2 Extracting Ontologies from XML Documents

Several strategies and tools were proposed for conversion of XML into OWL or RDF. A comparative analysis amongst those approaches was made in (Hacherouf, 2015). Bohring and Auer (2005) developed a xml2owl tool to generate an OWL ontology from XML schema by generating XSLT instances.

Reif et al. (2005) presented an approach "WEESA" for generating ontologies from an XML schema. It could be used to produce RDF metadata automatically from XML documents. Also from those tools are: XTR-RTO (Xu and Li, 2007), X2OWL (Ghawi and Cullot, 2009), the OWLMAP approach (Ferdinand et al., 2004), Janus (Bedini et al., 2010) (which has not been ready until now (Http://bivan.free.fr/Janus/,n.d.)]), JXML2OWL (Rodrigues et al., 2008), and XSPARQL (Bischof et al., 2011). However, the last two require the existence of the target ontology. Figure 10 depicts a small section of an XML document.

```
<? Xml version="1.0" encoding="UTF-8"?>
<Patient data card1>
<first-name> Smith </first-name>
<last-name> Bill </last-name>
<Gender> M </Gender>
<address> 23 elsalam street. </address>
<Birthdate> 18/4/1976 </birthdate>
<Report_date> June, 27 2010</Report_date>
<Blood group> A+ </Blood group>
<Blood Pressure> 150/80</Blood Pressure >
<Passport> t123456 </Passport>
<E-mail> eng.drf@g.com </E-mail>
<diagnosis> acute Q wave infarction - anteroseptal </diagnosis>
</ Patient data>
```

Figure 10. A small XML section.

4.3 Extracting Ontologies from Database Records

The transforming process of a relational database to the ontology has many transformation rules (Science, 2010). Those rules describe how the components of relational database (including rows, tables, constraints, columns, foreign keys, etc.,) can be translated into ontology segments (including classes, properties, axioms, and instances). Among those software tools: RDBToOnto ("RDBToOnto," n.d.), DB2OWL (Cullot, 2007), RDOTE (Vavliakis et al., 2013), R2O, D2RQ ("D2RQ Accessing Relational Databases as Virtual RDF Graphs," n.d.), KAON2 (KAON2, n.d.), DB2OntoModule, DataMaster protégé plugin (DataMaster, n.d.), and Ontology Generator (RDB2On). Figure 11 shows a small-scale EER diagram of the MySQL Workbench 8.0 CE for a patient local database schema used in our experiment.

4.4 Extracting Ontologies from ADL Standards

The translations between standards and the ontology are performed using the transformation modules and platforms shown in Figure 11. The main objective of each platform is to support valid EHR sharing of health data by transforming EHR standard ADL format into OWL format. A small part of the definition of the family history openEHR-EHR-COMPOSITION archetype is shown in Figure 13.

5 MATERIALS AND METHODS

IN the subsequent section, we will manipulate some of the software platforms and methods used in our experiment to convert each input data into an ontology format.

5.1 LinkEHR Studio Modeling Tool

LinkEHR Studio (Ibime, n.d.) is a software-independent archetype modeling tool. We used it to generate OWL format from ADL format, as depicted in Figure 14. It is implemented under the Java Eclipse platform. LinkEHR utilizes archetypes as a way to accomplish standardization and semantic interoperability of clinical distributed data (Maldonado et al., 2012). It can define archetypes based on many standards as openEHR, HL7 CDA, CEN/ISO 13606, ASTM CCR, HL7 FHIR, and CDISC ODM. LinkEHR has two main modules, as follows:

1. Multi-reference model editor of archetypes
2. Integration archetype editor, which allows the mapping of archetypes with legacy EHR data and the generation of XQuery standardization programs

5.2 *Protégé and Protégé Plug-ins*

Protégé is an open-source Java-based ontology modeling system with an OWL plug-in for building intelligent systems (Protege, n.d.). Protégé is corroborated by a strong academic community and by the government. It is used in many different areas, such as e-commerce, biomedicine, and organizational modeling. Protégé provides a GUI to help in ontology editing, such as creation, modification, reasoning and debugging. In it, multiple plug-ins are available: XML2OWL, Snow Owl, OntoGraf, and OntoBase.

For automatically converting databases into ontologies, Mogotlane, and Fonou-Dombeu revealed in a comparative analysis that OntoBase ("OntoBase," n.d.) outperforms other Protégé plug-ins (Mogotlane and Dombeu, 2016). In future implementations, we will use OntoBase as a Protégé plug-in to convert a database into an ontology format. For converting XML datasheets into ontologies, we will use XML2OWL (XML2OWL, n.d.), a Protégé plug-in that allows mapping XML to an OWL ontology.

6 CONCLUSION

New healthcare systems are compound of expert systems and soft computing techniques to help doctors and Specialists in making the correct decision at the correct time. New intelligent healthcare techniques and systems give access to many heterogeneous data sources such as virtual databases and knowledge bases. Semantic interoperability interests in providing meaningful use and exchange clinical data between many heterogeneous systems. In this paper, we move toward implementing the first module of our proposed semantic interoperability framework in a structured

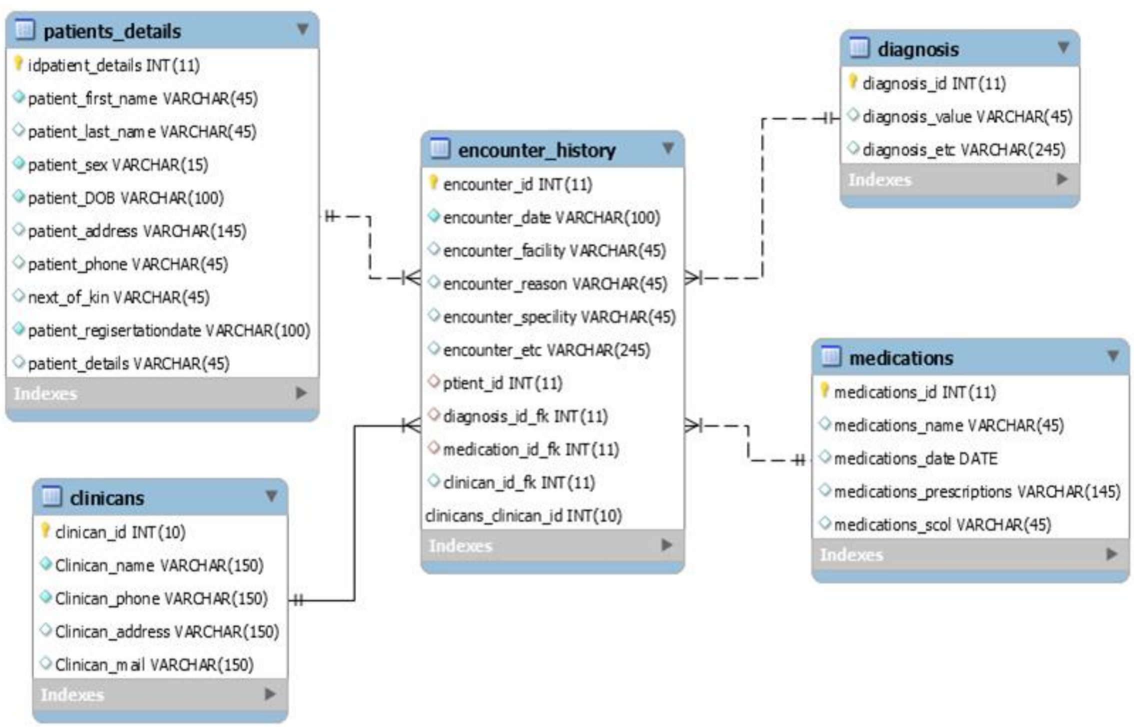


Figure 11. **A small-scale EER diagram of MySQL Workbench 8.0 CE for a local patient database.**

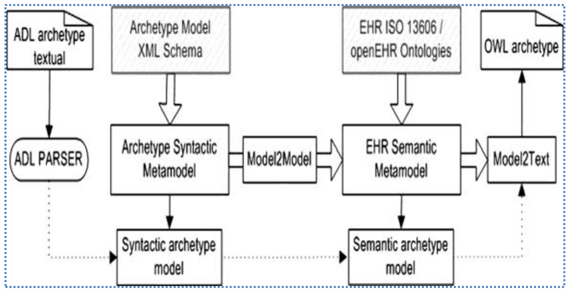


Figure 12. **The ADL archetypes in the OWL transforming process (Maldonado et al., 2012).**

heterogeneous EHR environment. Our main objective is to show how to exploit semantic web technologies to support EHR semantic interoperability. A unified framework based on a fuzzy ontology was proposed to solve the problem mentioned. Firstly; we tried to build an effective data-model to handle the integration issue of heterogeneous clinical systems. This research determines the suitable software and platforms needed in implementing the first module of the framework.

```
definition
  COMPOSITION[at0000] ∈ {   -- Family History
    category ∈ {
      DV_CODED_TEXT ∈ {
        defining_code ∈ {[openehr::433]}
      }
    }
    context ∈ {
      EVENT_CONTEXT ∈ {
        other_context ∈ {
          ITEM_TREE[at0003] ∈ {   -- Tree
            items cardinality ∈ {0..*; unordered} ∈ {
              allow_archetype ITEM[at0004] occurrences ∈ {0..*} ∈ {   -- Items
                include
                  archetype_id/value matches {/openEHR-EHR-CLUSTER\.document_entry_metadata(-[a-zA-Z0-9_]+)*\.v1/}
              }
            }
          }
        }
      }
    }
    content cardinality ∈ {1..*; unordered} ∈ {
      allow_archetype EVALUATION[at0001] occurrences ∈ {0..*} ∈ {   -- Family History
        include
          archetype_id/value matches {/openEHR-EHR-EVALUATION\.exclusion-family_history(-[a-zA-Z0-9_]+)*\.v1|openEHR-EHR-
EVALUATION\.family_history(-[a-zA-Z0-9_]+)*\.v1|openEHR-EHR-EVALUATION\.absence(-[a-zA-Z0-9_]+)*\.v1/}
      }
      allow_archetype EVALUATION[at0002] occurrences ∈ {0..*} ∈ {   -- Exclusion/Absence
        include
          archetype_id/value matches {/.*/}
      }
    }
  }
```

Figure 13. **An excerpt from the definition of family history in the openEHR-EHR-COMPOSITION archetype.**

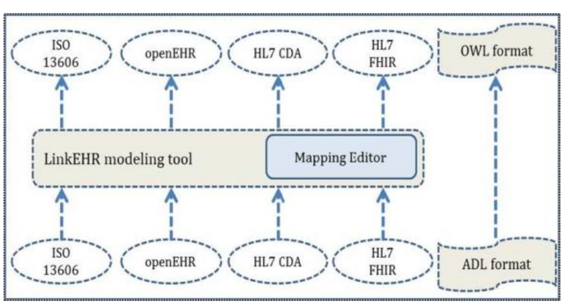


Figure 14. **LinkEHR is an independent reference archetype modeling tool.**

We expect the proposed approach to be highly intelligent and to have the ability to aggregate data with heterogeneous structures. It also provides physicians the ability to make uncertain and imprecise queries. The current limitation of this approach is that it is still under implementation. The next step will be practically measuring the degree of feasibility and applicability of the proposed algorithm in large healthcare record systems.

7 LIST OF ABBREVIATIONS

EHR	Electronic Health Record
IEEE	Institute of Electrical and Electronics Engineers
ASTM	American Society for Testing and Materials
ISO	International Organization for Standardization
HL7	Health Level 7
HL7 RIM	Health Level 7 Reference Information Model
ISO/TS	ISO technical specification
CEN/ISO	Comité European de Normalization/International Organization for Standardization
FHIR	Fast Healthcare Interoperability Resources
RDB	relational DataBase
XML	eXtensible Markup Language
W3C	World Wide Web Consortium
DTD	Document Type Definition
XSD	XML Schema Documentation
OWL	Web Ontology Language
RDF	Resource Description Framework
DBMS	DataBase Management System
ER	Entity Relationship
CSV	Comma Separated Values
CIMI	Clinical Information Modeling Initiative
CDISC	The Clinical Data Interchange Standards Consortium
DICOM	Digital Imaging and Communications in Medicine
IHE	Integrating the Healthcare Enterprise
DL	Description Logics
XSLT	Extensible Style Sheet Language Transformation
UML	Unified Medical Language
EER	enhanced entity–relationship

8 ACKNOWLEDGMENT

This work was partially supported by a National Research Foundation of Korea grant funded by the Korean Government (Ministry of Science and ICT) NRF-2017R1A2B2012337.

9 REFERENCES

Beale, T., & Heard, S. (Ocean I. (2007). Archetype Definitions and Principles. *Ocean Informatics*, (14 Mar 2007), 15. Retrieved from http://www.openehr.org/releases/1.0.2/architecture/am/archetype_principles.pdf

Bedini, I., Nguyen, B., & Gardarin, G. (2010). Janus: Automatic Ontology Builder from XSD files. *ArXiv Preprint ArXiv:1001.4892*, 2–3. Retrieved from http://arxiv.org/abs/1001.4892

Benson, T. (2016). *Principles of health interoperability : SNOMED CT, HL7, and FHIR*. [Cham]: Springer International Publishing AG, Switzerland.

Berges, I., Bermudez, J., & Illarramendi, A. (2012). Toward Semantic Interoperability of Electronic Health Records. *IEEE Transactions on Information Technology in Biomedicine*, *16*(3), 424–431. https://doi.org/10.1109/TITB.2011.2180917

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web: A New Form of Web Content That is Meaningful to Computers Will Unleash a Revolution of New Possibilities. *ScientificAmerican.Com*.

Bischof, S., Lopes, N., & Polleres, A. (2011). Improve efficiency of mapping data between XML and RDF with XSPARQL. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *6902 LNCS*, 232–237. https://doi.org/10.1007/978-3-642-23580-1_17

Blobel, B. (2002). *Analysis, design, and implementation of secure and interoperable distributed health information systems*. IOS Press.

Brewton, J., Yuan, X., & Akowuah, F. (2012). XML in Health Information Systems. *Proceedings of the International Conference on Bioinformatics & Computational Biology (BIOCOMP)*, (The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp)), 387–393.

Calvanese, D., Lenzerini, M., & Nardi, D. (1999). Unifying Class-Based Representation Formalisms. *CoRR*, *abs/1105.5*. Retrieved from http://arxiv.org/abs/1105.5452

Chandrasekaran, B., Josephson, J. R., & Benjamins, V. R. (n.d.). What Are Ontologies , and Why Do We Need Them ?

Cullot, N. (2007). DB2OWL : A tool for automatic database-to- ontology mapping. In *SEBD* (pp. 491–494).

D2RQ Accessing Relational Databases as Virtual RDF Graphs. (n.d.). Retrieved May 10, 2018, from http://d2rq.org/

Dadjoo, M., & Kheirkhah, E. (2015). An Approach For Transforming of Relational Databases to OWL Ontology, *6*(1).

DataMaster. (n.d.). DataMaster. Retrieved March 29, 2019, from https://protegewiki.stanford.edu/wiki/DataMaster

Desmond Mogotlane, K., & Fonou-Dombeu, J. V. (2016). Automatic Conversion of Relational Databases into Ontologies : A Comparative Analysis of Protege Plug-Ins Performances. *International Journal of Web & Semantic Technology*, *7*(3/4), 21–40. https://doi.org/10.5121/ijwest.2016.7403

Díaz, N., Cuéllar, M. P., Lilius, J., & Calvo-flores, M. D. (2014). Knowledge-Based Systems A fuzzy ontology for semantic modelling and recognition of human behaviour. https://doi.org/10.1016/j.knosys.2014.04.016

El Hajjamy, O., Alaoui, L., & Bahaj, M. (2018). Integration of Heterogeneous Classical Data Sources in an Ontological Database BT - Big Data, Cloud, and Applications. In Y. Tabii, M. Lazaar, M. Al Achhab, & N. Enneya (Eds.) (pp. 417–432). Cham: Springer International Publishing.

Elmogy, M., Adel, E., El-Sappagh, S., & Barakat, S. (2018). A unified fuzzy ontology for distributed electronic health record semantic interoperability (pp. 354–395).

Ferdinand, M., Zirpins, C., & Trastour, D. (2004). Lifting XML Schema to OWL. *Web Engineering. ICWE 2004. Lecture Notes in Computer Science*, (3140), 354–358. https://doi.org/10.1007/978-3-540-27834-4_44

Gerald Reif, H. G., and M. J. (2005). WEESA Web Engineering for Semantic Web Applications. *14th International Conference on World Wide Web*, 722–729. https://doi.org/10.1145/1060745.1060849

Ghawi, R., & Cullot, N., (2009). Building ontologies from XML data sources. *Proceedings - International Workshop on Database and Expert Systems Applications, DEXA*, 480–484. https://doi.org/10.1109/DEXA.2009.68

González, C., Blobel, B. G. M. E., & López, D. M. (2011). Ontology-based framework for electronic health records interoperability. *Studies in Health Technology and Informatics*, *169*(4), 694–698. https://doi.org/10.3233/978-1-60750-806-9-694

Hacherouf, M. (2015). Transforming XML documents to OWL ontologies: A survey. https://doi.org/10.1177/0165551514565972

Hannes Bohring and S¨oren Auer. (2005). Mapping XML to OWL Ontologies. *Leipziger Informatik-Tage*, *72*, 147--156.

Http://bivan.free.fr/Janus/. (n.d.). http://bivan.free.fr/Janus/. Retrieved November 1, 2018, from http://bivan.free.fr/Janus/

Ibime. (n.d.). Biomedical Informatics Group, Universidad Politécnica de Valencia. LinkEHR Normalization Platform. Retrieved December 31, 2018, from http://www.linkehr.com/

IEEE Standard Glossary of Software Engineering Terminology. (1990). *IEEE Std 610.12-1990*. IEEE. https://doi.org/doi: 10.1109/ieeestd.1990.101064

ISO reference definitions - guide 2 - 2004 - rev. (n.d.).

KAON2. (n.d.). KAON2 -- Ontology Management for the Semantic Web. Retrieved June 28, 2018, from http://kaon2.semanticweb.org/

Kiourtis, A., Mavrogiorgou, A., & Kyriazis, D. (2017). Aggregating Heterogeneous Health Data through an Ontological Common Health Language. In *2017 10th International Conference on Developments in eSystems Engineering (DeSE)* (pp. 175–181). https://doi.org/10.1109/DeSE.2017.9

Kubicek, H., Cimander, R., & Scholl, H. (2011). *Organizational Interoperability in E-Government - Lessons from 77 European Good-Practice Cases.*

Kuhanandha Mahalingam and Michael N.Huhns. 2000. "Ontology Tools For Semantic Reconciliation in Distributed Heterogeneous Information Environments." *Intelligent Automation & Soft Computing* 6 (3): 185–92. https://doi.org/10.1080/10798587.2000.10642787.

Kuck, G., (2004). Tim Berners-Lee's Semantic Web. *South African Journal of Information Management*, *6*(1).

Maldonado, J. A., Costa, C. M., Moner, D., Menárguez-Tortosa, M., Boscá, D., Miñarro Giménez, J. A., … Robles, M., (2012). Using the ResearchEHR platform to facilitate the practical application of the EHR standards. *Journal of Biomedical Informatics*, *45*(4), 746–762. https://doi.org/10.1016/j.jbi.2011.11.004

Martínez-Costa, C., Menárguez-Tortosa, M., Fernández-Breis, J. T., & Maldonado, J. A. (2009). A model-driven approach for representing clinical archetypes for Semantic Web environments. *Journal of Biomedical Informatics*, *42*(1), 150–164. https://doi.org/10.1016/j.jbi.2008.05.005

Martinez-Cruz, C., Blanco, I. J., & Vila, M. A. (2012). Ontologies versus relational databases: Are they so different? A comparison. *Artificial Intelligence Review*, *38*(4), 271–290. https://doi.org/10.1007/s10462-011-9251-9

Myłka, A., Myłka, A., Kryza, B., & Kitowski, J. (2012). Integration of heterogeneous data sources in an ontological knowledge base. *Computing and Informatics*, *31*(1), 189–223. https://doi.org/10.1080/17425255.2016.1211110

National Health Service - NHS England. (2015). Interoperability Handbook, 1–63. https://doi.org/10.1161/STROKEAHA.110.578583

OntoBase. (n.d.). Retrieved December 20, 2018, from https://code.google.com/archive/p/ontobase/

Ontological considerations in GIScience AU - Agarwal, P. (2005). *International Journal of Geographical Information Science*, *19*(5), 501–536. https://doi.org/10.1080/13658810500032321

Protege. (n.d.). Protege. Retrieved December 20, 2018, from https://protege.stanford.edu

Ramakrishnan, R., & Gehrke, J., (2009). Database Management System, 931. https://doi.org/10.1002/0471684228.egp03115

RDBToOnto. (n.d.). Retrieved December 31, 2018, from https://sourceforge.net/projects/rdbtoonto/

Rodrigues, T., Rosa, P., & Cardoso, J. (2008). Moving from syntactic to semantic organizations using JXML2OWL. *Computers in Industry*, *59*(8), 808–819. https://doi.org/10.1016/j.compind.2008.06.002

Schulz, S., & Martínez-Costa, C., (2013). How ontologies can improve semantic interoperability in health care. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *8268 LNAI*, 1–10. https://doi.org/10.1007/978-3-319-03916-9_1

Science, C., (2010). *Relational database as a source of ontology creation*. *Technology* (Vol. 1). IEEE.

Sheth, A. P., & Larson, J. A. (1990). Federated database systems for managing distributed, heterogeneous, and autonomous databases. *ACM Computing Surveys*, *22*(3), 183–236. https://doi.org/10.1145/96602.96604

Tapuria, A., Kalra, D., & Kobayashi, S. (2013). Contribution of clinical archetypes, and the challenges, towards achieving semantic interoperability for EHRs. *Healthcare Informatics Research*, *19*(4), 286–292. https://doi.org/10.4258/hir.2013.19.4.286

The Health Information Technology Policy Committee. (2015). Report to Congress: Challenges and Barriers to Interoperability, (December), 1–24.

Thi, P., Thuy, T., Lee, Y., & Lee, S. (2009). *DTD2OWL: automatic transforming XML documents into OWL ontology*. *Transformation*. ACM. https://doi.org/10.1145/1655925.1655949

Tho, Q. T., Hui, S. C., Member, S., Fong, A. C. M. & Member, S. (2006). Automatic Fuzzy Ontology Generation for Semantic Web, *18*(6), 842–856.

Vavliakis, K. N., Grollios, T. K., & Mitkas, P. A. (2013). RDOTE – Publishing Relational Databases into the Semantic Web The Journal of Systems and Software RDOTE – Publishing Relational Databases into the Semantic Web. *The Journal of Systems & Software*, *86*(1), 89–99. https://doi.org/10.1016/j.jss.2012.07.018

W3C. (n.d.). The XML data model. Retrieved September 29, 2018, from https://www.w3.org/XML/Datamodel.html

XML2OWL. (n.d.). XML2OWL. Retrieved December 20, 2018, from http://xml2owl.sourceforge.net/

Xu, J. X. J., & Li, W. L. W. (2007). Using Relational Database to Build OWL Ontology from XML Data Sources. *2007 International Conference on Computational Intelligence and Security Workshops (CISW 2007)*, (January), 124–127. https://doi.org/10.1109/CIS.Workshops.2007.139

Yahia, N., Mokhtar, S. A., & Ahmed, A. (2012). Automatic Generation of OWL Ontology from XML Data Source. *International Journal of Computer Science Issues*, *9*(2), 77–83.

10 NOTES ON CONTRIBUTORS

Ebtsam Adel is a computer technician at Mansoura University, Egypt. She had received her BSc and MSc degrees from the Faculty of Computers and Information, Mansoura University, Mansoura, Egypt. Now she is currently working on her Ph.D. degree at Information Systems Department, Faculty of Computers and Information, Mansoura University, Egypt. Her research interests involve computer vision and biomedical informatics.

Shaker El-Sappagh received the bachelor's degree in computer science and the master's degree from the Information Systems Department, Faculty of Computers and Information, Cairo University, Egypt, in 1997 and 2007, respectively, and the Ph.D. degree in computer science from the Information Systems Department, Faculty of Computers and Information, Mansura University, Mansura, Egypt, in 2015. In 2003, he joined the Department of Information Systems, Faculty of Computers and Information, Minia University, Egypt, as a Teaching Assistant. Since 2016, he has been with the Department of Information Systems, Faculty of Computers and Information, Benha University, as an Assistant Professor. He is currently a Post-Doctoral Fellow with the UWB Wireless Communications Research Center, Department of Information and Communication Engineering, Inha University, Incheon 22212, South Korea. He has publications in clinical decision support systems and semantic intelligence. His current research interests include machine learning, medical informatics, (fuzzy) ontology engineering, distributed and hybrid clinical decision support systems, semantic data modeling, fuzzy expert systems, and cloud computing. He is a reviewer for many journals, and he is very interested in disease diagnosis and treatment research.

Mohammed Elmogy is an associate professor at Information Technology Dept., Faculty of Computers and Information, Mansoura University, Egypt. He had received his B.Sc. and M.Sc. from Faculty of Engineering, Mansoura University, Mansoura, Egypt. He had received his Ph.D. from Informatics Department, MIN Faculty, Hamburg University, Hamburg, Germany in 2010. He is currently working as a visiting researcher at the bioengineering department, University of Louisville, Louisville, USA. He has authored/co-authored over 150 research publications in peer-reviewed reputed journals, book chapters, and conference proceedings. He has served as a reviewer for various international journals. He served as a technical program committee member in many workshops and conferences. He is a member of IEEE and ACM societies. He advised more than 25 master and doctoral graduates. His current research interests are computer vision, medical image analysis, machine learning, pattern recognition, and biomedical engineering.

Sherif Barakat is a professor at Information Systems Dept., Faculty of Computers and Information, Mansoura University, Egypt. He had received his B.Sc. and M.Sc. from Faculty of Science, Mansoura University, Mansoura, Egypt. He had received his Ph.D. from Faculty of Science, Helwan University, Helwan, Egypt in 2003. He has authored/co-authored over 50 research publications in peer-reviewed reputed journals, book chapters, and conference proceedings. His current research interests are machine learning, pattern recognition, and artificial intelligence.

Kyung-Sup Kwak received the Ph.D. degree from the University of California at San Diego in 1988. From 1988 to 1989, he was with Hughes Network Systems, San Diego, CA, USA. From 1989 to 1990, he was with the IBM Network Analysis Center, Research Triangle Park, NC, USA. Since then, he has been with the School of Information and Communication Engineering, Inha University, South Korea, as a Professor, where he had been the Dean of the Graduate School of Information Technology and Telecommunications from 2001 to 2002. He has been the Director of the UWB Wireless Communications Research Center (formerly Key National IT Research Center), South Korea, since 2003. In 2006, he served as the President of Korean Institute of Communication Sciences, and in 2009, the President of Korea Institute of Intelligent Transport Systems. In 2008, he had been selected for Inha Fellow Professor and now for Inha Hanlim Fellow Professor. Dr. Kwak published more than 200 peer-reviewed journal papers and served as TPC/Track chairs/organizing chairs for several IEEE related conferences. His research interests include wireless communications, UWB systems, sensor networks, WBAN, and nano communications. He was a recipient of the number of awards, including the Engineering College Achievement Award from Inha University, the LG Paper Award, the Motorola Paper Award, the Haedong Prize of research, and various government awards from the Ministry of ICT, the President, and the Prime Minister of Korea, for his excellent research performances.