Constructing Healthcare Ontologies of any Data Format

Athanasios Kiourtis, Argyro Mavrogiorgou, Dimosthenis Kyriazis
Department of Digital Systems
University of Piraeus
Piraeus, Greece
e-mail: {kiourtis, margy, dimos}@unipi.gr

***Abstract*—Current health systems are becoming stronger and more efficient, with promises for better and healthier life quality. However, one of the prerequisites for this promise is that health systems are able to connect and interact with each other, by quickly and seamlessly exchanging data, using open standards and bypassing interoperability constraints. For that purpose, several researches have been proposed focusing on the field of interoperability, dealing with specific one-to-one scenarios of data transformation. Among these solutions, the translation of healthcare data into ontologies is considered as a solution towards interoperability. However, healthcare data can be found in multiple formats, while most of the current approaches are dealing with specific data formats and designs. To address this gap, in this paper, we are proposing an extended approach to create an ontology model from any data format, taking into account complex cases arising from multiple data design styles, by transforming the healthcare data into XML Schema Graphs (XSG) for providing it as an input to the Apache Jena, and finally generating Resource Description Framework (RDF) entities.**

Keywords—healthcare; interoperability; ontologies; heterogeneous data

I. INTRODUCTION

Worldwide, healthcare is at the intersection of the introduction of disruptive digital innovations, with digital health revolutions playing a significant role in decreasing long-term healthcare costs, enabling better healthcare outcomes, and empowering both the patient and the healthcare provider with real-time data and connections with each other [1]. However, for this promise it is considered crucial that health systems are able to communicate with each other, by quickly and seamlessly exchanging data using open standards [2]. Nevertheless, one of the biggest issues in healthcare research is the ability of researchers to obtain healthcare record data, as information silos of healthcare data exist across both private and public sectors, affecting both patient care and medical research [3]. In order to overcome these information silos, there is a great need for legislative directives that allow and encourage data sharing across both federal and commercial healthcare environments, including security requirements to protect personally identifiable information and protected health information [4]. Even with these directives, issues will remain regarding information exchange, since data has to be shared with as many stakeholders as possible. To this domain, interoperability is the only sustainable way for letting systems to talk with one another and getting the complete image of a patient [5], considered of great importance to overcome today's fragmented and proprietary global health systems, allowing stakeholders to make the most of health data, including extending, upgrading, and preserving it. According to [6][7], while global annual health spending reached $7.077 trillion dollars in 2015, this metric should increase to $8.734 trillion dollars by 2020. Moreover, according to [8], researchers predict that the healthcare market will have a value of $18.7 billion by 2020, up from a size of $5.8 billion in 2015. What is more, the global healthcare middleware market is expected to reach $3.07 billion by 2023 from $1.90 billion in 2018, in order to overcome healthcare interoperability issues [9]. Such issues include exchanging healthcare data between disparate systems to support care coordination, population health management initiatives and most importantly to stay Health Insurance Portability and Accountability (HIPAA) compliant to support meaningful use of data and protect the privacy of individuals [10]. Based on [11], 86% of the mistakes made in healthcare are due to administrative reasons, where 30% of clinical tests have to be re-ordered since the results cannot be found. Moreover, the same report has showed that patient charts cannot be found on 30% of visits, while about 80% of all serious medical errors involve miscommunication during care transitions.

Henceforth, it is undeniable that healthcare provisioning and research require healthcare data to be restructured into a common format and standard terminologies. In general, healthcare data can be either structured such as in relational databases, or semi-structured such as in eXtensible Markup Language (XML) data sources [12]. Therefore, data sources can be heterogeneous in syntax, schema, or semantics, thus making data interoperation a difficult task [13]. Consequently, syntactic heterogeneity is caused by the use of different models or languages, schematic heterogeneity is caused by structural differences, while semantic heterogeneity is caused by different meanings or interpretations of the data in various contexts. In all these cases, interoperability can be achieved by manufacturing medical domain ontologies for representing the different concepts, relationships, and axioms among the different healthcare datasets, and for representing medical terminology systems [14]. More specifically, ontologies provide a promised technology to solve especially the semantic heterogeneity problem, as a formal, explicit

 ISBN: 978-1-61208-688-0

specification of a shared conceptualization [15], referring to an abstract model of phenomena in the world, having identified the relevant concepts of those phenomena. ‘Explicit’ means that the type of concepts used and the constraints on their use are explicitly defined. ‘Formal’ refers to the fact that the ontology should be machine readable, while ‘Shared’ means that the ontology should capture consensual knowledge accepted by the communities.

In all the aforementioned cases, healthcare interoperability is currently delivered through the HL7 Fast Healthcare Interoperability (HL7 FHIR) standard [16], which despite its wide adoption, it still needs much time to become a global healthcare data exchange standard, as there exist systems that still produce data that are not related to it. Taking into consideration these challenges, in [17] a holistic approach has been presented for achieving interoperability through the transformation of healthcare data into its corresponding HL7 FHIR structure. Shortly, the provided mechanism was building the healthcare ontologies that were primarily stored into a triplestore, in order to identify and compare their syntactic and semantic similarity with the HL7 FHIR Resources. Consequently, according to the aggregation of the syntactic and semantic similarity results, the matching and translation to the HL7 FHIR was taking place. In this paper, we are going to describe the mechanism of the automatic ontology creation of any data format that is currently in [17], taking into consideration that dozens of ontology creation mechanisms have been developed over the last decade, and nearly all of them are able to perform transformations only on specific data formats and standards (e.g., csv or xml files).

The remainder of the paper is organized as follows. In Section II, the related work is illustrated, while Section III depicts the overall approach with regards to the automatic ontology creation of any data format. Section IV includes the results of the evaluation of the proposed approach, while Section V presents our concluding remarks and future directions.

II. Related Work

A. Ontologies in Healthcare

Current healthcare systems need to have the ability to communicate complex and detailed medical data in a secure, automated, and effective way. This can be achieved by building medical domain ontologies [18], in order to represent medical terminology systems. In general, an ontology represents a common, shareable, and reusable view of a particular application domain that gives meaning to information structures that are exchanged among different information systems. As a result, ontology’s aim is to achieve shareable knowledge for transmitting it between different stakeholders and application systems. Henceforth, the main goal is to create semantics in a generic way, providing the basis for agreement within a domain. Generally, the advantage of wrapping each information source to a local ontology is to allow the development of a source ontology independently of other sources or ontologies. Engineering new ontologies is not a deterministic process, since many design decisions must be made, while the designers’ backgrounds and the target application, influence their decisions in different ways.

In the context of healthcare, ontologies may bring the indispensable integration of knowledge and data [19]. They are composed of existing medical knowledge, providing a way for building systems that help healthcare providers to perform clinical actions in a more efficient and effective way. Ontologies also specify semantic-based criteria that support different statistical aggregations for different purposes in healthcare, improving the accuracy of diagnoses by providing real time correlations of symptoms, test results and individual medical histories through standards-based systems for systematic crosschecking diagnoses. To this context, the authors in [20] developed a healthcare information system based on ontology methods, where healthcare practitioners were able to construct ontologies about new diseases, being easier to predict upcoming diseases. Furthermore, the authors in [21] built an ontology-based healthcare context information model to implement a ubiquitous environment. In that case, contextual information was extracted and classified to implement the healthcare services using the context information model that could be defined through ontologies. As a result, application and healthcare service developers could use the sensed information in various environments by authoring device- and space-specific ontologies based on this common ontology. The research of [22] should be mentioned, where the authors proposed a framework and toolset that could provide a secure single point of access to a client’s full picture of her personal health information, by proposing an ontology-based framework. This framework was an independent tool that could automatically gather and combine a client’s health information from the various providers in their circle of care and provide the information securely and electronically without inconveniencing the client with multiple requests and sharing agreements. Finally, the researchers in [23] developed an ontology-based toolkit for improving the field of semantic interoperability in healthcare data. Shortly, the authors identified and specified the potential data sources, they conceptualized their semantic meaning and defined to what extent routine data could be used as a measure of the process or outcome of care required, in order to finally formalize and validate the final ontology.

B. Ontology creation from different data sources

In the field of ontology creation, several strategies and researchers have been developed for deriving ontologies mainly from heterogeneous XML data sources. Some of these approaches target on a general mapping between XML and Resource Description Framework (RDF) [24], while others aim at mapping XML Schema [25] to Web Ontology Language (OWL) [26]. In this domain, the authors in [27] proposed direct mappings from XML Schema to OWL, describing mappings from XML to RDF graphs, where the generated instances did not necessarily respect the ontology created from the XML Schema. The XML Schema to OWL mapping process was based on a set of interpretation and transformation rules from XML Schema to OWL. What is more, the authors in [28] developed a tool to create an OWL ontology from an XML Schema, and convert XML data to OWL instances compliant to the created ontology, through the development of four Extensible Style Sheet Language Transformation (XSLT) [29] instances to transform XML files to OWL, without any other intervention on semantics and structures during the transformation. The research in [30] proposed a tool that aimed at building an OWL ontology from an XML data source. This method was based on XML Schema to automatically generate the ontology structure, as well as a set of mapping bridges between the entities of the XML data source and the created ontology, contributing into query translation between OWL and XML. The authors in [31] proposed a framework that aimed to generate an ontology from a large source of XML Schemas. They presented a set of patterns that enabled the direct, and automatic transformation from XML Schema into OWL, allowing the integration of huge amounts of XML data in the Semantic Web. They focused on an advanced logical representation of XML Schema components and presented an implementation that was possible to mine XML Schema sources in order to extract enough knowledge to build semantically correct ontologies with considerable expressivity. Moreover, the research in [32] proposed an approach to construct OWL ontology from XML document with the help of entity relation model. The authors proposed an XML-to-Relational (XTR) mapping approach to map an XML document to an entity-relation model, and then a Relational-to-Ontology (RTO) mapping approach to map an entity-relation model to an OWL ontology. What is more, the authors in [33] proposed an approach to integrate heterogeneous XML sources using an ontology-based mediation architecture. The ontology integration process contained the schema transformation and ontology merging steps, used for modelling each XML source as a local RDF ontology. Finally, the authors in [34] were based on XML Schema to build an ontology. In the case that the schema did not exist, it could be automatically generated from the source XML document, coping with all the possible complex cases that could arise from different XML Schema design styles.

C. *Key contributions of the proposed approach*

Several researches have been proposed so far in the literature focusing on the healthcare interoperability field, whereas dealing with specific one-to-one scenarios of data transformation. Among these solutions, the translation of healthcare data into ontologies is considered as a solution towards interoperability. However, healthcare data can be found in multiple data formats, while during ontology transformations, different terms are produced for the same concept. Most of the developed approaches are dealing with specific data format (i.e., mostly XML data format), while they do not tackle the question on how to create the ontology model, if there is not any XML Schema available. To address this gap, in this paper we are proposing an extended approach to create an ontology model from any data format, taking into account even the most complex cases arising from different data formats and design styles.

III. PROPOSED APPROACH

This study is based on an existing approach [17] that can transform any healthcare dataset of any format and nature, into HL7 FHIR through the translation of the latter into ontologies, and their matching through syntactic and semantic similarities. In our case, we are going to propose a mechanism for describing the automatic creation of the healthcare ontologies for the ingested healthcare dataset. Fig. 1 illustrates the overall architecture of the mechanism as depicted in [17], where the components that are highlighted with grey color are going to be ignored since they have already been discussed. Therefore, in the proposed mechanism only the Ontology Building System is going to be described in deep details.

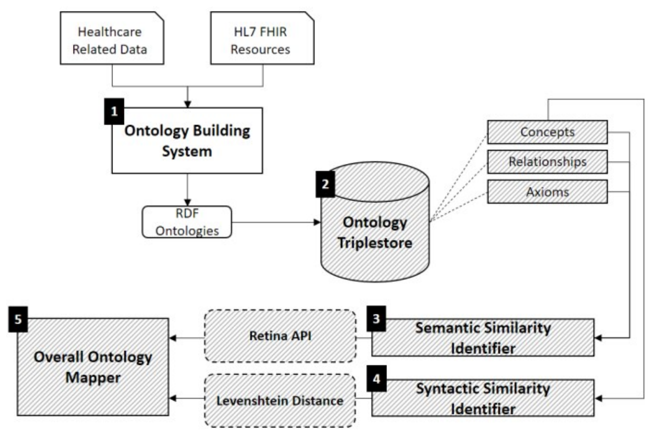

Figure 1. Overall architecture.

A. *Ontology Building System*

The objective of the Ontology Building System is to gather the XML Schema of the gathered healthcare datasets, and conclude in building the corresponding ontologies as described in the following five steps (Fig. 2):

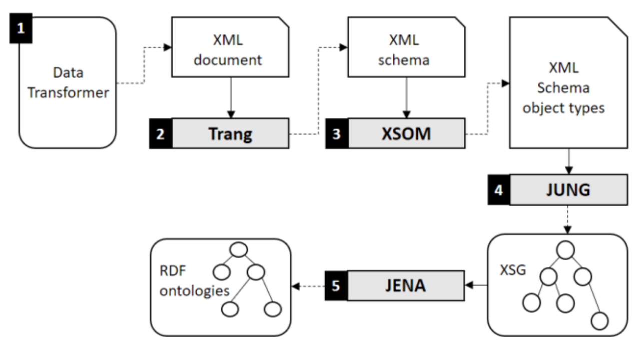

Figure 2. Ontology Building System.

1. Each healthcare dataset is being preprocessed through a document parser (Data Transformer) in order to identify the type of the document in which the dataset is stored (e.g., JSON, CSV, etc.). In any case, the Data Transformer converts the type of the healthcare dataset into XML format, by identifying the different elements of the document, as well as the parent and the child nodes, along

 ISBN: 978-1-61208-688-0

with their attributes. Henceforth, all the converted data is being stored into an XML document.

2. The created XML document is transformed into XML Schema using the Trang API [35]. Shortly, Trang performs conversions between different schema languages for XML, and is constructed around a RELAX NG object model [36] designed to support schema conversion, taking as input a file written in XML syntax and producing an XML Schema.

3. Afterwards, the XML Schema is analyzed using the XML Schema Object Model (XSOM) [37]. The SOM provides a set of classes in the System.Xml.Schema namespace that allows the reading of a schema from a file or the creation of a schema in-memory. The schema can then be traversed, edited, compiled, validated, or written to a fileXSOM. In this context, XSOM allows applications to easily parse XML Schema documents and inspect information into them.

4. Sequentially, the output of the XSOM is used as an input for the Java Universal Network/Graph framework (JUNG) [38], which is used for graph-based manipulations. The latter, generates an XML Schema Graph (XSG) [39] that describes the schema in the same way whatever its design style is.

5. Finally, the XSG is used as an input to the Apache Jena [40], in order to generate ontologies in the form of RDF entities [41]. It should be mentioned that RDF entities emerge from complex types, element group declarations, and attribute-group declarations according to specific matching rules. Apart from these, ontologies contain different object properties that emerge from relationships that are of type element-subelement relationships, while they contain datatype properties that emerge from attributes and from simple types.

IV. EVALUATION

A. *Use Case description*

The use case exploits a healthcare dataset structured in CSV format, covering all the previous steps. The dataset that has been used to evaluate the efficiency of the proposed approach (Fig. 3) was a sub-dataset of anonymized citizens' personal information, derived from Karolinska Institute [42].

```
subject;gender;dateOfBirth;dateOfDeath;causeOfDeath
ID0054619;1;1/6/1954;NA;NA
ID1012175;0;1/6/1968;NA;NA
ID0332172;0;1/11/1961;NA;NA
ID0186244;0;1/7/1963;NA;NA
ID0782623;1;1/5/1957;NA;NA
ID0944842;1;1/5/1961;NA;NA
ID0131218;1;1/7/1976;NA;NA
ID0131218;1;1/12/1951;NA;NA
ID0911095;1;1/11/1945;NA;NA
ID0951273;0;1/5/1928;20060212;C2299
ID0695153;1;1/4/1980;NA;NA
ID0695153;0;1/4/1977;NA;NA
ID0695153;1;1/4/1959;NA;NA
ID0695153;1;1/6/1951;NA;NA
ID1164801;0;1/1/1978;NA;NA
ID1409311;0;1/9/1960;NA;NA
ID0097871;0;1/8/1950;NA;NA
ID1095873;1;1/6/1924;NA;NA
ID0250381;0;1/10/1954;NA;NA
```

Figure 3. Use case dataset sample.

In more details, it consists of 5000 different instances of the personal information of certain citizens about their:
(i) subject: string that depicts the personal identifier
(ii) gender: integer that varies between 0 (male) and 1 female
(iii) date of birth: date/time that depicts the birth date
(iv) date of death: date/time that depicts the death date
(v) cause of death: string that depicts the cause of death

B. *Application of the Ontology Building System*

In order to depict the proposed ontology creation process, after providing as an input the described dataset, we gather the following results, for each different step of the described process of Section III.

Initially, after the preprocessing of the CSV use case dataset through the Data Transformer, the results of the XML creation of this dataset are represented in Fig. 4.

```xml
<root>
    <entry id="1">
        <subject>ID0054619</subject>
        <gender>1</gender>
        <dateOfBirth>1/6/1954</dateOfBirth>
        <dateOfDeath>NA</dateOfDeath>
        <causeOfDeath>NA</causeOfDeath>
    </entry>
    <entry id="2">
        <subject>ID1012175</subject>
        <gender>0</gender>
        <dateOfBirth>1/6/1968</dateOfBirth>
        <dateOfDeath>NA</dateOfDeath>
        <causeOfDeath>NA</causeOfDeath>
    </entry>
    <entry id="3">
        <subject>ID0332172</subject>
        <gender>0</gender>
        <dateOfBirth>1/11/1961</dateOfBirth>
        <dateOfDeath>NA</dateOfDeath>
        <causeOfDeath>NA</causeOfDeath>
    </entry>
    <entry id="4">
        <subject>ID0186244</subject>
        <gender>0</gender>
        <dateOfBirth>1/7/1963</dateOfBirth>
        <dateOfDeath>NA</dateOfDeath>
        <causeOfDeath>NA</causeOfDeath>
    </entry>
```

Figure 4. XML document of the use case dataset.

In sequence, the XML Schema that is created through the derived XML document is represented in Fig. 5.

```xml
<?xml version="1.0" encoding="UTF-8"?>
<xsd:schema targetNamespace=""
elementFormDefault="qualified"
xmlns:xsd="http://www.w3.org/2001/XMLSchema">
    <xsd:element name="root">
        <xsd:complexType mixed="true">
            <xsd:sequence>
                <xsd:element name="entry" maxOccurs="unbounded">
                    <xsd:complexType mixed="true">
                        <xsd:sequence>
                            <xsd:element name="subject" minOccurs="0"
                            type="xsd:normalizedString"/>
                            <xsd:element name="gender" minOccurs="0"
                            type="xsd:int"/>
                            <xsd:element name="dateOfBirth" minOccurs="0"
                            type="xsd:normalizedString"/>
                            <xsd:element name="dateOfDeath" minOccurs="0"
                            type="xsd:normalizedString"/>
                            <xsd:element name="causeOfDeath" minOccurs="0"
                            type="xsd:normalizedString"/>
                        </xsd:sequence>
                        <xsd:attribute name="id" type="xsd:int"
                        use="required"/>
                    </xsd:complexType>
                </xsd:element>
            </xsd:sequence>
        </xsd:complexType>
    </xsd:element>
</xsd:schema>
```

Figure 5. XML Schema of the use case dataset.

The next step of the proposed mechanism deals with the analysis of the XML Schema through the XSOM, whose results are provided as an input to the JUNG in order to finally generate the XSG. The latter is represented in Fig. 6.

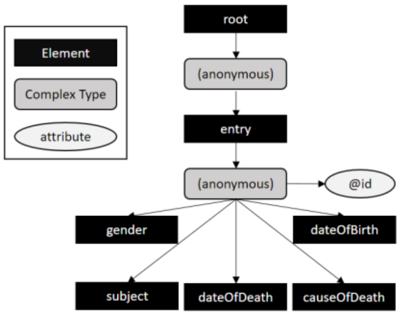

Figure 6. XSG of the use case dataset.

To this end, the XSG is used as an input to the Apache Jena to generate the RDF entities that are depicted in Fig. 7.

```
<rdf:Description rdf:nodeID="A14">
  <owl:allValuesFrom
  rdf:resource="http://uri-not-absolute.com#entry"/>
  <owl:onProperty rdf:resource="http://uri-not-absolute.com#entry"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Restriction"/>
</rdf:Description>
<rdf:Description rdf:nodeID="A6">
  <owl:allValuesFrom
  rdf:resource="http://www.w3.org/2001/XMLSchema#normalizedString"/>
  <owl:onProperty rdf:resource="http://uri-not-absolute.com#subject"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Restriction"/>
</rdf:Description>
<rdf:Description rdf:nodeID="A11">
  <owl:allValuesFrom
  rdf:resource="http://www.w3.org/2001/XMLSchema#normalizedString"/>
  <owl:onProperty rdf:resource="http://uri-not-absolute.com#dateOfBirth"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Restriction"/>
</rdf:Description>
<rdf:Description rdf:nodeID="A0">
  <owl:maxCardinality
  rdf:datatype="http://www.w3.org/2001/XMLSchema#int">1</owl:maxCardinality>
  <owl:onProperty rdf:resource="http://uri-not-absolute.com#dateOfBirth"/>
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#Restriction"/>
</rdf:Description>
```

Figure 7. RDF entities of the use case dataset.

Finally, Fig. 8 represents the final output of the mechanism, depicting the ontological hierarchical tree of the use case dataset, as it is visualized by WebVOWL [43].

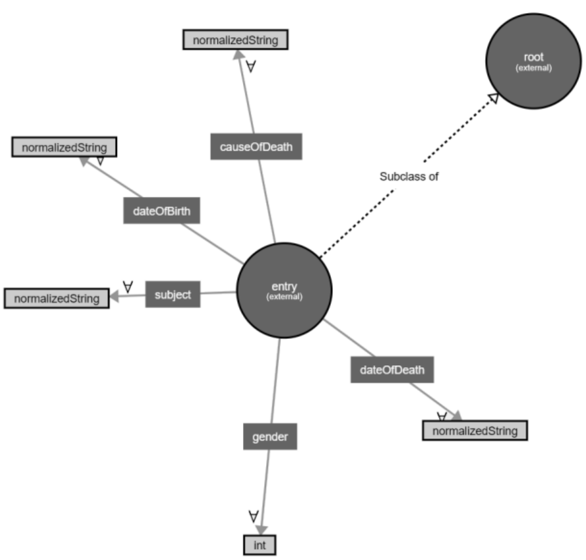

Figure 8. Ontological hierarchical tree of the use case dataset.

V. CONCLUSIONS

Currently, health systems are becoming stronger and more efficient, with promises for better and healthier life quality. However, one of the prerequisites for this promise is that health systems are able to interact with each other, by seamlessly exchanging data using open standards, thus bypassing interoperability constraints. In this domain, healthcare ontologies have been developed to resolve data heterogeneity issues, resulting in interoperability by knowledge sharing and reuse. Most of the developed approaches that aim in ontology building are dealing with specific data format, while they do not tackle the question on how to create the ontology model, if there is not any XML Schema available. In order to address the gap of automated ontology creation, a previous research was considered about a healthcare ontology matching mechanism that has the ability of transforming healthcare data to HL7 FHIR format, by building healthcare ontologies, and finding syntactic and semantic similarities among these ontologies and the ontologies of the HL7 FHIR Resources. In that case, an extended approach was presented that is able to create an ontology model from any data format, taking into account complex cases arising from different data formats and design styles. According to the results of the evaluation of the proposed approach, the overall ontology creation mechanism provided reliable results that are compatible with the manually derived results.

Generally, writing an automated ontology building mechanism that is able to cover multiple data formats is still a challenging research task, since there exist several mechanisms that provide the same functionalities. In this context, we will work on the evaluation of the Ontology Building System with the similar mechanisms that were studied and are currently used, in order to identify potential errors and gaps that could be addressed in the current approach. What is more, we want to perform a similar comparison and evaluation concerning the other sub-mechanisms of the current approach for healthcare interoperability, in order to conclude to more reliable and efficient results. To this end, we will continue on the evaluation of the Ontology Building Systems with datasets of different sizes, standards and formats, respecting privacy issues, based on the mechanism developed in [44].

ACKNOWLEDGMENT

The CrowdHEALTH project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 727560.

REFERENCES

[1] Is digital health the future of healthcare?, https://www.forbes.com/sites/quora/2018/04/04/is-digital-health-the-future-of-healthcare/#201d7f592556 [last accessed: 07/01/2019]

[2] Data Integration and Interoperability: ISO/OGC, https://www.directionsmag.com/article/3396 [last accessed: 07/01/2019]

[3] Why Health Systems Must Embrace Interoperability Now, https://www.webpt.com/blog/post/so-long-to-silos-why-

health-systems-must-embrace-interoperability-now [last accessed: 07/01/2019]

[4] Data Silos, https://www.himss.org/news/data-silos-together-we-can-bust-them [last accessed: 07/01/2019]

[5] Moving Forward with Interoperability, https://www.philips.com/a-w/about/news/archive/blogs/innovation-matters/moving-forward-with-interoperability-how-patients-will-drive-change.html [last accessed: 07/01/2019]

[6] The Economist Intelligence Unit (EIU) is the research and analysis division of The Economist Group and the world leader in global business intelligence, http://www.eiu.com/home.aspx [last accessed: 07/01/2019]

[7] Life Sciences and Healthcare, https://www2.deloitte.com/global/en/pages/life-sciences-and-healthcare/articles/global-health-care-sector-outlook.html [last accessed: 07/01/2019]

[8] Healthcare analytics market worth $18.7 billion by 2020, https://www.whatech.com/market-research/medical/archive/192126-healthcare-analytics-market-worth-18-7-billion-by-2020 [last accessed: 07/01/2019]

[9] Healthy Growth Forecast for Medical Devices Global Market, https://www.bccresearch.com/pressroom/hlc/healthy-growth-forecast-for-medical-devices-global-market [last accessed: 07/01/2019]

[10] Healthcare Security: Understanding HIPAA Compliance and its Role in Patient Data Protection, https://digitalguardian.com/blog/healthcare-security-understanding-hipaa-compliance-and-its-role-patient-data-protection [last accessed: 07/01/2019]

[11] 3 Signs You Should Invest in Healthcare Automation, https://www.formstack.com/blog/2017/invest-in-healthcare-automation/ [last accessed: 07/01/2019]

[12] R. Arnon, L. Seligman, and R. Costello, "XML, Databases, and Interoperability." Federal Database Colloquium, AFCEA, San Diego, 1999.

[13] I. F. Cruz, and X. Huiyong, "The role of ontologies in data integration." Engineering intelligent systems for electrical engineering and communications, vol. 13(4), p.245, 2005.

[14] B. A. Gylys, and M. E. Wedding, “Medical terminology systems: a body systems approach”, FA Davis 2017.

[15] E. M. Gontier, “Web Semantic and Ontology”, 2015.

[16] HL7 FHIR, https://www.hl7.org/fhir/ [last accessed: 07/01/2019]

[17] A. Kiourtis, A. Mavrogiorgou, D. Kyriazis, and S. Nifakos, “Aggregating the Syntactic and Semantic Similarity of Healthcare Data towards their Transformation to HL7 FHIR through Ontology Matching”, Int. Journal of Medical Informatics (IJMI), 2018. under review

[18] P. Wennerberg, “Aligning medical domain ontologies for clinical query extraction”, 12th conference of the EU chapter of the association for computational linguistics: student research workshop. Association for Comp. Linguistics, 2009.

[19] Ontologies in healthcare supporting the integration of knowledge, http://www.dataversity.net/ontologies-in-healthcare-supporting-the-integration-of-knowledge/ [last accessed: 07/01/2019]

[20] V. Kumar, “Ontology based public healthcare system in Internet of Things (IoT)”, Procedia Computer Science, vol. 50, pp. 99-102, 2015.

[21] J. Kim, and K. Y. Chung, “Ontology-based healthcare context information model to implement ubiquitous environment”, Multimedia Tools and Applications, vol. 71(2), pp. 873-888, 2014.

[22] M. Azarm, and L. Peyton, “An ontology for a patient-centric healthcare interoperability framework”, IEEE/ACM International Workshop on Software Engineering in Healthcare Systems (SEHS), pp. 34-41, 2018.

[23] H. Liyanage, P. Krause, and S. de Lusignan, “Using ontologies to improve semantic interoperability in health data”. Journal of innovation in health informatics, vol. 22(2), pp. 309-315, 2015.

[24] Resource Description Framework (RDF), https://www.w3.org/RDF/ [last accessed: 07/01/2019]

[25] Xml schema, https://www.w3.org/standards/xml/schema [last accessed: 07/01/2019]

[26] Web Ontology Language (OWL), https://www.w3.org/OWL/ [last accessed: 07/01/2019]

[27] M. Ferdinand, C. Zirpins, and D. Trastour, “Lifting XML Schema to OWL”, In Web Engineering - 4th International Conference, Munich, Germany, pp. 354– 358, 2004.

[28] H. Bohring, and S. Auer, “Mapping XML to OWL Ontologies”, In Leipziger Informatik-Tage, vol. 72, pp. 147–156. Society, Washington, DC, USA, 2007.

[29] Extensible Stylesheet Lagnuage (XSLT), https://www.w3schools.com/xml/xsl_intro.asp [last accessed: 07/01/2019]

[30] R. Ghawi, and N. Cullot, “Building Ontologies from XML Data Sources”, 1st Workshop on Modelling & Visualization of XML and Semantic Web Data, pp. 480-484, 2009

[31] I. Bedini, G. Gardarin, and B. Nguyen, “Transforming XML Schema to OWL Using Patterns”, 5th IEEE Int. Conference on Semantic Computing, Palo Alto: US, pp.1-8, 2011.

[32] J. Xu, and W. Li, “Using Relational Database to Build OWL Ontology from XML Data Sources”, International Conference on Computational Intelligence and Security Workshops, pp. 124– 127, 2007.

[33] I. F. Cruz, X. Huiyong, and H. Feihong, “An ontology-based framework for XML semantic integration”, pp. 217-226, 2004.

[34] N. Yahia, S. A. Mokhtar, and A. Ahmed, “Automatic generation of OWL ontology from XML data source”, arXiv preprint arXiv:1206.0570, 2012.

[35] Trang, http://www.thaiopensource.com/relaxng/trang.html [last accessed: 07/01/2019]

[36] Relax NG, http://www.kohsuke.org/relaxng/ [last accessed: 07/01/2019]

[37] XML Schema Object Model (XSOM), https://directory.fsf.org/wiki/XSOM [last accessed: 07/01/2019]

[38] Jung, http://jung.sourceforge.net/ [last accessed: 07/01/2019]

[39] M. Hacherouf, S. N. Bahloul, and C. Cruz, “Transforming XML documents to OWL ontologies: A survey”, Journal of Information Science, vol. 41(2), pp. 242-259, 2015.

[40] Apache Jena, https://jena.apache.org/ [last accessed: 07/01/2019]

[41] RDF entities, https://www.infowebml.ws/rdf-owl/ [last accessed: 07/01/2019]

[42] Karolinska Institute, https://ki.se/ [last accessed: 07/01/2019]

[43] Web Ontology Visualizer, http://vowl.visualdataweb.org/webvowl.html [last accessed: 07/01/2019]

[44] A. Kiourtis, A, Mavrogiorgou, and D. Kyriazis, “Towards a Secure Semantic Knowledge of Healthcare Data Through Structural Ontological Transformations”, In Joint Conference on Knowledge-Based Software Engineering, pp. 178-188, 2018.