Semantic interoperability challenges to process large amount of data perspectives in forensic and legal medicine

Marie-Christine Jaulent, Damien Leprovost, Jean Charlet, Remy Choquet

▶ To cite this version:

HAL Id: hal-01431816

https://hal.sorbonne-universite.fr/hal-01431816

Submitted on 11 Jan 2017

Semantic Interoperability Challenges to process Large Amount of Data Perspectives in Forensic and Legal Medicine

Marie-Christine Jaulent[a], Damien Leprovost[a], Jean Charlet[a], Remy Choquet[a,b]

[a] *INSERM, U1142, LIMICS, F-75006, Paris, France; Sorbonne Universités, UPMC Université Paris 06, UMR_S 1142, LIMICS, F-75006, Paris, France; Université Paris 13, Sorbonne Paris Cité, UMR_S 1142, LIMICS, F-93430, Villetaneuse, France*
[b] *BNDMR, Assistance Publique Hôpitaux de Paris, Hôpital Necker Enfants Malades, Paris, France*

Abstract

Although the Big Data approach seems promising in various analytic uses, sharing or integrating data within the same analysis space remains a complex task as existing data is highly heterogeneous and difficult to compare. In this position paper, we address the Variety and Veracity dimensions of Big Data when integrating, sharing and reusing large amount of heterogeneous data for data analysis and decision making applications in the healthcare domain. Many issues are raised by the necessity to conform Big Data to standards in order to make data more interoperable both by humans or computations such as data mining. In this paper, we discuss how ontologies (computerized meaning) can contribute to the improvement of information sharing and address the problem of data sharing together with semantic interoperability data frameworks. We then introduce the main steps required for building domain ontologies as they could be implemented in the context of Forensic and Legal medicine. We conclude with a particular emphasis on the current limitations in healthcare data standardization and the importance of knowledge formalisation. for the coming years, in order to maximise data re-use in forensic and legal medicine.

Introduction

The exponential production of data in our societies and the exponential interest in exploiting these data led to a new area known as "big data" [1]. Following the Gartner Research's definition, "*big data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making*" [2]. Since 2011, data storage and data analysis have been the first challenges of this field. In particular, new data processing techniques, novel multivariate techniques have to be conceived to handle increasingly higher dimensional datasets. For instance, nonlinear dimensionality reduction (NLDR) techniques were applied to analyse complex and big data in order to improve the estimation of age in living persons in Forensic science [3]. NLDR is preferred to Principal Components Analysis (PCA) techniques, which was usually the common method to deal with these multi-variable analysis, due to the explosion of complexity inherits to the Big Data context. This is one of the numerous examples of the changes needed to face this new generation of data [4].

As the field is evolving, there is not a unique and common perception of Big Data. In the healthcare domain, a definition has been provided by the IMIA (*International Medical Informatics Association*) working group on "Data Mining and Big Data Analytics": "*Big Data are data whose scale, diversity, and complexity require new architecture,*

techniques, algorithms, and analytics to manage it and extract value and hidden knowledge from it" [5, 6]. In the keynote of the 2014 Yearbook of IMIA dedicated on "*Big Data and Smart Health Strategies*" [7], Bellazzi et al. precise for the medical domain the usual characterization of big Data by four Vs [8]: "*Volume means "large" in size, Variety is related to the different format in which the data are collected, from structured information to text, from images to signals, Velocity is related to the need of processing data at a pace that is fast enough to support decision-making and Veracity is related to the uncertain nature that is often connected to data of large volumes, very often without pre-processing or collected with limited quality control*". A direct consequence of the presence of one or more of these factors is to change the technological architecture used to manage the data and the algorithms designed to analyse them.

An important barrier for electronic information exchanges is the lack of interoperability (see the glossary for a definition of syntactic and semantic interoperability) between information systems especially when data is scattered within different information systems, used for different purposes, are stored in different technical systems or formats and represent different meanings. To address those issues, technical frameworks accessing and comparing distributed and heterogeneous data with consistent semantics are needed to enable cooperative research and progress in the comprehension of data exchanged.

In this position paper, we introduce a semantic interoperability framework to address the Variety and Veracity dimensions when integrating, sharing and reusing heterogeneous data for data analysis and decision making applications. We will illustrate our proposition by delineating the challenge of bridging together data from siloed information systems. We then discuss current bottlenecks in setting interoperability solutions and we argue on the importance of ontologies (see the glossary for a definition of ontology) and semantics applied to data driven processes and we finally discuss the building of an ontology in Forensic and Legal medicine.

Challenges in accessing, sharing and reusing massive heterogeneous data

Big data are receiving an increasing attention in many sectors such as biomedicine and healthcare, facing massive heterogeneous data coming from various sources that could help in decision-making or public health policy making. Such objectives require to overcome obstacles that arise from the different ways data are collected and coded into healthcare systems and information systems.

The different potential sources for Big Data analysis in healthcare include electronic health records (primary care, hospitals, prescription data), population-based registries, genomics, (multi)-national disease registries, health insurance databases, and newer sources of patient data and information from home monitoring, wearables, mobile apps, online patient forums, and social media. Moreover, though health care is collecting large amounts of data, it is also important to collect data from environmental conditions. The environment's contribution to health has been conceptualized as the exposome, defined as every exposure to which an individual is subjected from conception to death, requiring consideration of the nature of the exposures and their changes and can be considered as general external (urban environment, climate factors, social capital, stress) [9]. As these data increasingly become available and accessible in standard way, there is a growing interest in sharing and reusing such data for many uses and studies.

Lately, there has been an intense focus directed to the issue of developing and maintaining shareable, multipurpose, high-quality computable algorithms operating on different data sources thanks to standardized access to data [10]. Recognizing the

growing need of data integration and its relation to sustainability and scalability issues, the FAIR Data Initiative has been launched in January 2014 [11]. This new movement provides guidelines to adopt standards in order to develop FAIR data repositories and networks where valuable scientific data is 'FAIR' in the sense of being Findable, Accessible, Interoperable and Reusable [12]. An important step in the FAIR Data approach is to publish existing and new datasets in a semantically interoperable format that can be understood by other humans (not the producers of the data) and computer systems in specific settings. By semantically annotating data items to be integrated with metadata, computer systems can be used to (semi-automatically) combine different data sources, resulting in richer knowledge discovery activities. The table 1 shows the why and how of the different challenges.

Table1 : Challenges in accessing, sharing and reusing massive heterogeneous data

Challenges	Why ?	How ?
#1 Access data	Be accessible for data means being understandable. But many formats and data encodings are available and used to store data on different systems.	Data should be described in standard and formal models that are robust to data evolution.
#2 Share data	As when accessing, sharing poses the problem of understandings. Shared data must be manageable by all future system. This challenge is a prerequisite to the next challenge: reuse.	Data provider or transformer must adopt standards in order to produce findable, accessible, interoperable and reusable data (FAIR [12] and OpenData [38])
#3 Reuse data	Published data has been extracted in a specific context. Understanding these data is related to an understanding of this context, otherwise misinterpretations can occur	The data condition and context have to be exported as metadata and linked to the data itself.

This initiative aims at facilitating the necessary breaking down of information systems but the challenge of exploiting data in legacy systems still remains as well as issues of exploiting new kinds of data for which standards do not yet exist (Internet of Things, Social Media such as discussion boards, Twitter, etc.). If healthcare studies now focus these new sources of data [13, 14], the lack of base reference and the unstructured and moving nature of the common language do this field a remaining challenge for automatic comprehension [15].

Since the last decade of the past century considerable efforts have been invested into the development of standards for health information representation such as the recent open Clinical Information Modeling Initiative (http://www.opencimi.org), with an increasing focus on semantic interoperability. CIMI is an international collaboration working on a common format for specifying the representation of health information content so that semantically interoperable information may be created and shared in health records, messages and documents. CIMI provide these specifications in different formats, such as the Archetype Definition Language (ADL) from the openEHR Foundation (ISO 13606.2) and the Unified Modeling Language (UML) from the Object Management Group (OMG) with the intent that the users of these specifications can convert them into their local formats. Semantic interoperability (SIOp) is the ability of a system to record and communicate information across system boundaries, with a

minimum of information loss. Standardised and consensual vocabularies, terminologies, data dictionaries, data models, value sets and conceptual information models are also facilitators of SIOp, opposed to local coding systems, local information models or natural language. Insufficient SIOp bears the risk of suboptimal decision-making.

In the rest of the paper, we introduce the rationale of SIOp by emphasizing the role of ontologies to represent the studied domain semantics and we describe a method to build such ontology in the forensic and legal medicine domain that could help in analysing vast amounts of generated data within the forensic digital domain which represents one required building block to achieve data sharing or analysis at larger scale.

Semantic interoperability frameworks

Nowadays, SIOp is a broadly used paradigm to approach the problem of sharing data from heterogeneous data sources [16]. System interoperability has been identified as a key challenge, critical to success. It is now well established that SIOp relies on the adoption of interoperability standards (reference information models/templates, value sets and terminologies) that support information sharing among systems [17]. One difficulty is that despite efforts from Standards Development Organizations (SDOs) (Health Level Seven International (HL7), Digital Imaging and Communications in Medicine (DICOM) or CEN Technical Committee 251 (CEN TC251)) and regardless of the international initiative of "Integrating the Healthcare Enterprise" (IHE), most clinical data in Electronic Healthcare Records (EHR) applications are still not natively interoperable.

Although the development of operational platforms faces the standardisation issues, several semantic interoperability platforms were nevertheless developed in various academic research projects (mostly European) to apply this paradigm and bring a proof of concept for data analysis by querying heterogeneous and distributed data sources. Examples are given by the EHR4CR European project [18] and the DebugIT European project [19]. The rationale of such systems is to define the intended meaning (semantics) of the existing data to ensure coherent interpretation by humans and processing tools, even when data are not coded the same way [20]. Such approach requires huge efforts from experts to reach a consensus about the semantics of the domain and the possible standard vocabularies to be used to bridge data.

To alleviate the burden of the elicitation of semantics, it is widely accepted today that the use of formal representations of semantics (formalized in ontologies) and associated semantic web techniques are promising tools to improve the quality of the processing chain of medical data. The semantic interpretation of the data stored in databases ensures semantic interoperability in the data sharing process. Some existing frameworks are based on ontologies to unify structural models and terminologies, together with relevant mapping sets. This approach has been tested in the context of the EU Framework Program 7 DebugIT project [19] and TRANSFoRm project [21]. Figure 1 shows the three level architecture required to set an semantic interoperability framework.
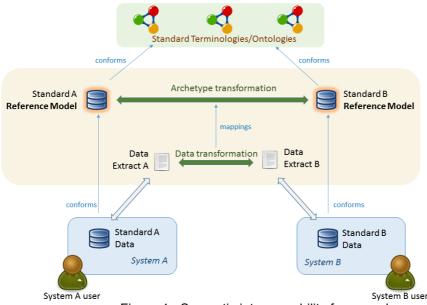
Figure 1 : Semantic interoperability framework

An illustration of the use of ontologies and SIOp frameworks is the European DebugIT project where many ontologies were used. First, a domain specific ontology was designed in order to enable correct domain reasoning within decision making process at a later stage (the studied domain was the evolution of antibiotic resistance). Second, ontologies to describe data sources were set; a specific quality measure framework was also set in order to annotate source data with quality metrics [23, 20]. Then, a set of semantic web tools were used to integrate data pulled from diverse sources within the domain ontology (as instances) so data could be compared at European level [24]. The SIOp defined in DebugIT is as generic as possible since data ontologies and domain ontologies could be re-used in another domains. The domain ontology was then used as a pivot model to both integrate heterogeneous data sources as well as to build decision support applications within the frame of antibiotic resistance studies. Implementing such framework in the forensic and legal medicine is a realistic solution but will require formalising semantics of the field through a domain ontology at first.

Building an ontology in forensic and legal medicine

A wide-accepted definition of an ontology is proposed by T. Gruber [22], "shared specification of a conceptualization". In this definition, ontology defines mainly the concepts and relationships that are relevant for modelling a domain and that will be used by knowledge-based systems.

Forensic medicine is the science that deals with the application of medical knowledge to legal questions (use of medical testimony in law cases). Many sources of information are classically taken into consideration for decision making, such as clinical, dental, radiological or pathology. One important tool of forensic medicine is the autopsy, usually conducted to determine the cause of death but also for identification of the dead. For dead or living persons, forensic DNA databases are now well established in many countries in the world. For example, in France the Fichier National Automatisé des Empreintes Génétiques (FNAEG) has been set up in 1998. In the USA, the FBI has organized the Combined DNA Index System (CODIS) database and software [25]. Gathering Local, State and National DNA Index Systems (LDIS, SDIS and NDIS) and originally intended for sex offenders, they have since then been extended to include almost any criminal offender or arrested person [26]. The challenge for building a SIOp

in this field is to set an ontology that will be used to describe the data sources just mentioned and probably others. The remaining of the article presents a method to build an ontology from the available textual sources in the field (literature, guidelines, documentation, medical records, pathology reports, etc.).

A consequence of an accepted definition of an ontology is that knowledge can be captured and made available to both machines and humans: we need a methodology which interacts with humans – i.e. clinician expert – in order to build in use knowledge. Thus, we propose a methodology focusing on natural language based on the facts that (i) domain stakeholder writes texts to describe situations and (ii) these texts are evidences of an implicit modelization. The methodology follows methodological guidelines that were first outlined in [27] mainly by Bachimont and consists in four steps, 1) choice of the corpus, 2) semantic normalization of the terms introduced in the ontology, 3) formalization of the meaning of the knowledge primitives and 4) operationalization using knowledge representation languages.

- The first step consists in building a corpus of texts in order to allow Natural Language Processing tools to perform good analysis of the terminology domain. For the purpose of forensic and legal medicine, it will be necessary to identify the source of these texts (different kinds of forensic or legal reports).

- The goal of the second step is to reach a semantic agreement about the meaning of the labels used for naming the concepts. Meaning of labels, extracted from the first step, is specified clearly using a relevant semantic theory. In some previous work, we built a differential ontology where terms are turned into notions based on differential semantics [28]. Practically, similarities and differences of each notion with respect to its neighbors have to be expressed: its parent-notion and its siblings-notions. The result is taxonomy of notions, where the meaning of a node is given by the gathering of all similarities and differences attached to the notions found from the root notion (the more generic) to the node in question.

- The ontological tree obtained in the preceding step allows to disambiguate the notions and to clarify their meanings for a domain-specific application. The transition to extensional semantics aims at linking the notions to a set of referents. The notions become concepts behaving as formal primitives and being part of a referential ontology. Each concept refers to a set of objects in the domain (its extension).

- The fourth and last step of the methodology allows enrich referential concepts with possible computational operations in formal logics: it is the computational ontology. The system uses an operational knowledge representation language that allows particular inferences. For a language based on description logics (like OWL), these inferences are mainly subsumption tests and classification.

This Methodology was used in few domains as pneumology [28] or emergency medicine [30]. The built domain ontology may be linked to referential ontologies as FMA in order to share anatomical concepts [31] or Upper-level ontology as Dolce [32] or OBO [33] in order to support ontology mapping and improve interoperability. Finally, this process requires medical expert only at particular validation times. Natural Language Processing tools can ease the process by starting from professional's texts and thus saving the expert's time. Advantages concern the fact that, in texts, terms design not only concepts but also concepts in use and that possibly relevant implicit knowledge is not lost in intermediate interpretations.

Discussion and Conclusion

Although the promises of Big Data to manage large amount of generated data within an operational setting is emerging and in some cases is already set as a standard (sensor data, social networks, genetic data, etc.), the use of Big Data as an integration paradigm for decision support is still limited. Within the healthcare domain, Big Data methods have trouble in comparing data without having a proper implicit knowledge (or models) of the processed data formalised. For instance, considering data in EHRs (Electronic Healthcare Records), the complexity of the information artefacts used to describe data (SNOMED-CT nomenclature is worth more than 350 000 clinical terms, UMLS is medical meta-thesaurus more than 1,5 million terms) makes the task difficult [34]. It is also widely accepted that terminologies and ontologies cannot capture all implicit knowledge so the data can be fully interpretable, it must be associated with information models or contextual models (statistics). Ontologies can also be associated with rule based systems that can more efficiently reflect contextual information such as conditional assertions (*if the result of the antiobiogram is positive, then the patient needs a specific therapy).* Ontologies will provide you with the formal definitions of biological exams, antiobiograms, positive result, therapy, and you will be able to assert that this therapy can be used in many cases. But is this individual situation, the reason of the specific therapy, when you will provide it and how, cannot be captured in an ontology. The existence of some implicit knowledge that cannot be formalized is nowadays a research topic debated in conferences on context [35]

Nowadays, Semantic Interoperability platforms are key components when data needs to be exchanged within operational or decisional processes within a domain. However, in the healthcare domain, operational solutions for SIOp are still confidential due to the inability of health applications to conform to interoperability standards. We indicate here some reasons that have been in particular highlighted in the context of the European ASSESS-CT project [36]: *(i) applications tend to operate in silos, for specific set of operations, don't share or combine resources; (ii) a lack of well accepted standards and good practices hampers the handling and use of personal data, and there is no accepted model and standard to be legally valid; (iii) patients have not been well enough represented in the development of codes of practice on the handling of health data for research, or in the definition of standards for capturing and communicating health records and research data sets; (iv) there is a lack of awareness among communities about the health information and knowledge representation interoperability standards that exist in the areas of clinical research and clinical care, and still a limited availability of mapping tools that can support the semantic harmonisation of heterogeneous health and life sciences data; and (v) even when operational data is re-used, every data integration or Big Data projects establishes their own framework of policies and standard operating rules, appoints an ethics and governance board, and spends time developing materials and gaining approvals from data providers, patient representatives and other stakeholders.*

In this paper, we argue that ontologies can be used to enrich data and to query data stored in large heterogeneous databases [37]. When properly modelled, they can support decision making when integrating within a specific operational process. Several proofs of concepts have been done in various contexts [18, 19, 24]. The main obstacle come from the amount of expert involvement and the needed time to build ontologies as well as the necessity to specify very clearly the domain to be formalized. There are today

very efficient methods to build ontologies using large amount of expertise embedded in textual sources and we believe that the Forensic and Legal medicine will benefit from designing such an ontology.

Commented glossary

Knowledge organization system (KOS)

Language allowing to represent knowledge, with or without formal constraints. Lexicon, thesaurus, classification, terminology, ontology are different KOS.

Ontology

An ontology is a "shared specification of a conceptualization". In this definition of T. Gruber, ontology defines formally the concepts and relationships that are relevant for modelling a domain. Because it is described in a formal language (Cf. infra) , ontology is used by knowledge-based systems. In this context, thesaurus or ICD-10 classification are not an ontologies because not described in formal language.

Formal language

In mathematics, computer science, and linguistics, a formal language is a set of strings of symbols that may be constrained by rules that are specific to it. This property allows to represent axiomatic systems and to maintain truth in formula. Formal languages are the base of reasoning systems, and consequently symbolic Artificial Intelligence.

Formal semantics

Formal semantics is semantics of a resource or a system described with formal language. Because formal semantics of human language is very criticized, many authors attribute formal semantics only to computer. Different semantics are associated to human being in relation with their language capability (e.g. generative or interpretative semantics - https://en.wikipedia.org/wiki/Generative_semantics)

Syntactic interoperability

It refers to the ability of a system to exchange information with another system with respect of standard languages (for description as for transport).

Semantic interoperability

It refers to the ability of a system to exchange data with another system and to understand descriptions in exactly the same manner. Semantic interoperability is necessarily based on Ontology. Syntactic interoperability is necessary to a semantic interoperability implementation.

Bibliography

[1] Google. Google Trends for Big Data, 2013.

[2] http://www-01.ibm.com/software/data/bigdata/

[3] Lefèvre T., Chariot P., Chauvin P. Multivariate methods for the analysis of complex and big data in forensic sciences. Application to age estimation in living persons. *Forensic Science International*, Forensic Sci Int. 2016 Sep;266:581.

[4] Qiang Yang. 2013. Big data, lifelong machine learning and transfer learning. In Proceedings of the sixth ACM international conference on Web search and data mining (WSDM '13). ACM, New York, NY, USA, 505-506.

[5] Smitha T, V. Suresh Kumar, Applications of big data in data mining, International Journal of Emerging Technology and Advanced Engineering, 7, 3, 2013 (www.ijetae.com).

[6] Peek N, Sun J, Holmes J, Martin-Sanchez F, Bellazzi R, Biomedical and Healthcare Analytics on Big Data, AMIA 2013 Symposium Proceedings, 2013, November, 1116-1117.

[7] Christoph U. Lehmann, Brigitte Séroussi, Marie-Christine Jaulent. Big3. Editorial. *Yearb Med Inform.* 2014; 9(1): 6–7. Published online 2014 May 22. doi: 10.15265/IY-2014-0030

[8] http://www.schattauer.de/en/magazine/subject-areas/journals-a-z/imia-yearbook/imia-yearbook-2014/issue/1973/issue/special/manuscript/21157/show.html

[9] Martin Sanchez F, Gray K, Bellazzi R, Lopez-Campos G. Exposome informatics: considerations for the design of future biomedical research information systems. J Am Med Inform Assoc. 2014 May-Jun;21(3):386-90

[10] Weng C, Tu SW, Sim I, Richesson R.. Formal representation of eligibility criteria: a literature review. J Biomed Inform 2010;43:451-67

[11] http://www.datafairport.org

[12] Wilkinson M. D., Dumontier M, Aalbersberg I et al. The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3, Article number: 160018 (2016)

[13] Michael J Paul and Mark Dredze. You are what you tweet: Analyzing twitter for public health. In ICWSM, pages 265–272, 2011.

[14] Alex Lamb, Michael J Paul, and Mark Dredze. Separating fact from fear: Tracking flu infections on twitter. In HLT-NAACL, pages 789–795, 2013.

[15] Hassan Saif, Miriam Fernandez, Yulan He, and Harith Alani. Senticircles for contextual and conceptual semantic sentiment analysis of twitter. In The Semantic Web: Trends and Challenges, pages 83–98. Springer, 2014.

[16] Moreno-Conde, A., Moner, D., da Cruz, W.D., Santos, M.R., Maldonado, J.A., Robles, M., Kalra, D. Clinical Information modeling processes for semantic interoperability of electronic health records: systematic review and inductive analysis. J. Am. Med. Inform. Assoc. 2015; Jul;22(4):925-34

[17] Do NV, Barnhill R, Heermann-Do KA, Salzman KL, Gimbel RW. The military health system's personal health record pilot with Microsoft Health-Vault and Google Health. J Am Med Inform Assoc 2011;18:118–24.

[18] de Moor GD, Sundgren M, Kalra D, Schmidt A, Dugas M, Claerhout B, Karakoyun T, Ohmann C, Lastic PY, Ammour N, Kush R, Dupont D, Cuggia M, Daniel C, Thienpont G, Coorevits P. Using Electronic Health Records for Clinical Research: the Case of the EHR4CR Project. J Biomed Inform. 2015 Feb;53:162-73.

[19] Lovis C, Colaert D, Stroetmann VN. DebugIT for patient safety - improving the treatment with antibiotics through multimedia data mining of heterogeneous clinical data. Stud Health Technol Inform. 2008;136:641-6

[20] Assélé Kama A, Choquet R, Mels G, Daniel C, Charlet J, Jaulent MC. An ontological approach for the exploitation of clinical data. Stud Health Technol Inform. 2013;192:142-6.

[21] Ethier JF, Dameron O, Curcin V, McGilchrist M, Verheij R,Arvanitis T,Taweel A, Delaney B, Burgun A. A unified structural/terminological interoperability framework based on LexEVS: application to TRANSFoRm. J Am Med Inform Assoc 2013;20:986–994.

[22] T. Gruber. Toward Principles for the Design of Ontologies Used for Knowledge Sharing. International Journal Human-Computer Studies 43, p.907-928, 1995

[23] Choquet R, Qouiyd S, Ouagne D, Pasche E, Daniel C, Boussaid O, Jaulent MC. The Information Quality Triangle: a methodology to assess clinical information quality. *Stud Health Technol Inform.* 2010;160(Pt 1):699-703.

[24] Schober D, Boeker M, Bullenkamp J, Huszka C, Depraetere K, Teodoro D, Nadah N, Choquet R, Daniel C, Schulz S. The DebugIT core ontology: semantic integration of antibiotics resistance patterns. Stud Health Technol Inform. 2010;160(Pt 2):1060-4.

[25] https://www.fbi.gov/about-us/lab/biometric-analysis/codis/codis-and-ndis-fact-sheet

[26] Roth, Andrea. "Maryland v. King and the Wonderful, Horrible DNA Revolution in Law Enforcement." Ohio St. J. Crim. L. 11 (2013): 295.

[27] B. Bachimont, A. Isaac, R. Troncy, Semantic Commitment for Designing Ontologies : A proposal. *European Knowledge Acquisition Workshop*, (2002)

[28] F. Rastier, M. Cavazza, A. Abeille, *Sémantique pour l'analyse.* Masson, Paris, (1994)

[29] [A Baneyx, J. Charlet, and MC Jaulent, Methodology to Build Medical Ontology from Textual Resources, AMIA Annu Symp Proc. 2006; 2006: 21–25.

[30] J Charlet, L Mazuel, G Declerck, P Miroux, And P Gayet. Describing Localized Diseases in Medical Ontology: An FMA-based Algorithm. Stud Health Technol Inform. 2014;205:1023-7.

[31] MEJINO J. V., AGONCILLO A. V., RICKARD K. L. & ROSSE C. (2003). Representing complexity in part-whole relationships within the Foundational Model of Anatomy. AMIA ... Annual Symposium proceedings / AMIA Symposium. AMIA Symposium, p. 450–4.

[32] Gangemi, A., Guarino, N., Masolo, C., Oltramari, A. and Schneider, L., Sweetening Ontologies with DOLCE, Proceedings of EKAW, 2002

[33] T Bittner, B. Smith. Formal ontologies for space and time (2003)

[34] Martínez-Costa C, Kalra D, Schulz S. Improving EHR semantic interoperability: future vision and challenges. Stud Health Technol Inform. 2014;205:589-93.

[35] P Brézillon. Context in Artificial Intelligence: I. A survey of the literature. Computers and artificial intelligence 18, 1999, 321-340

[36] http://assess-ct.eu/start0.html

[37] Möller, R., Neuenstadt, C., Özçep, Ö. L., & Wandelt, S. (2013). Advances in Accessing Big Data with Expressive Ontologies. In I. J. Timm & M. Thimm (Eds.), KI 2013: Advances in Artificial Intelligence: 36th Annual German Conference on AI, Koblenz, Germany, September 16-20, 2013. Proceedings (pp. 118–129). Berlin, Heidelberg: Springer Berlin Heidelberg. http://doi.org/10.1007/978-3-642-40942-4_11

[38] C. Bizer, T. Heath, T. Berners-Lee. Linked Data—The Story So Far. International Journal on Semantic Web and Information Systems. 5 (3): 1–22. 2009.