SEMANTIC WEB TECHNOLOGIES FOR DATA ORGANIZING AND PROCESSING IN E-BUSINESS PROCESSES

Ramona Markoska[1], Nikola Rendevski[1], Ilija Hristoski[2], Tome Dimovski[1]

[1]Faculty of ICT, UKLO-Bitola; [2]Faculty of Economics, UKLO-Bitola, Macedonia

Abstract: Today's business processes have a mandatory online component as an added value. As a part of the processes of continuous qualitative adjustments and improvements, there is a need for special treatment of different categories of information, collected from the online work environment. In e-business processes, the information is simultaneously a raw material, but also a final product. In a dynamic online environment, information of interest for processing is constantly changing. The added value in the processing of information is of great importance, and therefore new ICT tools and solutions are needed, which can better respond to this challenge. Therefore, the modern trends in the management of ICT solutions in the e-business processes are in line with the semantic web standards, and are focused on the collection, organization and processing of information of importance. This paper analyzes the benefits of using of semantic web technologies, such as ontologies in the definition of semantically compatible data formats and semantic NoSQL databases for the organization and storage of semantically organized data. The presented practical example describes the processes of organizing semantic data and their placement in the MongoDB database and explains the immediate use of semantic web technologies to manipulate data in e-business.

Keywords: Management of ICT, Semantic WEB, e-business

1. INTRODUCTION

The main idea of the semantic web, according the vision of the director of the Consortium W3C, Tim Berners-Lee, is to create online contents, understandable to machines, which means to the servers and network computers. Unlike that, the former concept of the web was realized to be much close with the logic and cognitive processes of the humans [2]. Further, analysing today's e-business processes in a variety of business activities, it comes to the conclusion that, in many different countries, e-commerce is fastest growing e-business part over the last decade. In order to well satisfy the needs of today's e-business working environment, e-commerce is supported by many various technologies, such as: semantic web, data mining, and neural networks [1]. From the other side, a good organization of information is crucial, in order to achieve a competitive advantage in the e-commerce processes. [1],[3],[4]. For this purpose, there are defined pattern discovery methods for online data: data clustering, data filtering, user's path analysis, and association rules in order to find searching association pattern. In the classic web, structured according to the human's logic, the following kind of problems appears: *Information symmetry problem*- which appears when different websites show different and multiple inappropriate information, for the same search

query; and *Search problem-* which refers to the situation when the search for some product did not shown the proper relevant usable result according product's generic attributes [3]. The semantic web technologies help overcome these problems, and the technological groundwork for their broad realization has been implemented during the activities of data organizing and processing in e-commerce. Through a generalized analysis of these two problems, it come to the semantic web technologies, in which various tools are synergistically approaching in finding solutions, through specific ways of data manipulation. Furthermore, it leads to the creation and interconnection of complete semantic tools, for defining specific data formats, supported by appropriate databases and search techniques.

2. SEMANTIC WEB TECHNOLOGIES' DESCRIPTION

As an improvement of the classic web of documents, W3C is helping to build a technology stack to support a web of data, where semantic data stored in databases is used. The ultimate goal of the web of data is to enable computers to do more useful work and to develop systems that can support trusted interactions over the network. The term "semantic web" refers to W3C's vision of the web of linked data. The semantic web technologies enable people to create data stores on the web, build vocabularies, and write rules for handling data [5]. From the point of view of the technologies used in its development, semantic web expanding qualitatively the classical hypertext network, without replacing it. In a constantly growing, globally unregulated medium, such as WWW, the main goal of semantic web technologies is to solve the two previously described problems, through the following procedures:

- *Creating standardized ontologies for different areas of e-business*, more precisely, for e-commerce, which means a creation of specific data formats, according the agreed general rules, in order to have a universal description of the products and their properties. Starting from the general definition of ontology as the formal naming and defining of the types, attributes and relationships of entities that are common parts in a particular domain of discourse, [2] one of the most effective approaches is a creation of ontologies through the transformation of existing international standards, such as eclass, EDIFACT, ebXML. [4]

- *Reduction of information asymmetry in e-business/e-commerce,* to a model of perfect information, in line with market requirements. Bearing in mind that anyone can publish information about anything in the semantic network, mechanisms for guaranteeing the authenticity of the data and their authorization are needed [4]. Perfect information in the above context, means a balance and a link between the information needs of users-consumers and the intentions for promotion and advertisement of companies-manufacturers, over WWW.

2.1. SEMANTIC WEB STACK

In order to understand semantic web technologies it is necessary to consider the structure of the semantic web stack. The semantic web stack displays the hierarchy of the languages and technologies used. Obviously, the technologies from the upper layers are being developed and upgraded to the technologies used in the layers below, in the shown hierarchy.

During the evolution of the standards and technologies of the semantic network, a significant moment is the promotion of Resource Description Framework (RDF), and Resource Framework Description Schema (RDFS). These two metadata standards and easy web publishing schemes, developed in 1999, played a significant role in establishing semantic web standards and the development of the previous semantic web stack in 2000, which is revised in 2018. Figure 1. shows the semantic web stack, updated in 2018, where the current intersection of the development of semantic web technologies is presented.
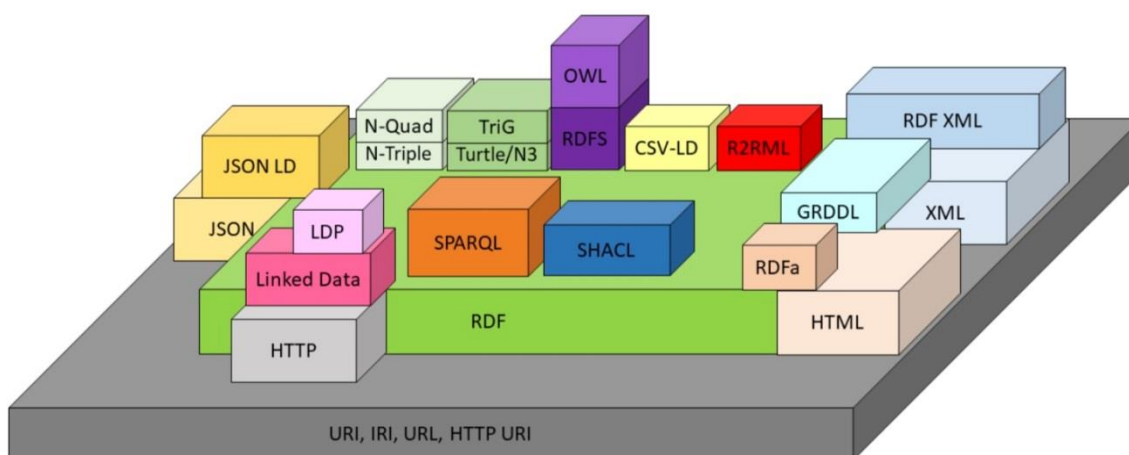

Figure 1. A new version of the semantic web stack 2018 [9]

Analysing the new semantic web stack, it can be noticed that there is no strict hierarchy in layers as in the previous version. Because many of the technologies have more functionalities that in practice can replace or complement some other existing technologies, flexibility and creativity are ensured when working, even when strict rules are not followed. Below is a brief description of the more important technologies and the way they are interconnected. [3] [5][6][9].

The bottom layer consists of a set of standard protocols for initial identification and communication, through the web. This layer provides a *mechanism for identifying certain resources and parts of namespaces* (URI, IRI, URL, HTTP URI), client-server application protocols (request-response), and create conditions for initial communication for distributed information environments (HTTP)[5] [6].

The second layer consist of a set of tools, languages and protocols [6] [7], interconnected and complemented in a flexible whole, such as:

- *Extensible Markup Language (XML)*, which is used to describe two different types of data- human and machine readable, and represents a work base for upgrading with other languages such as RDF and OWL.

- *Resource Description Framework (RDF)*, is a standard model for data interchange on the Web. RDF extends the linking structure of the Web to use URIs to name the relationship between things as well as the two ends of the link, which this is usually referred to as a "triplet".

- *JavaScript Object Notation (JSON)* is an open-standard file format that uses human-readable text to transmit data objects consisting of attribute–value pairs and array data types, which is a very often used for asynchronous browser–server

551

communication, including as a replacement for XML., and

- Well known from before, *Hypertext Markup Language (HTML),* the standard markup language for creating web pages and web applications, and *Hypertext Transfer Protocol (HTTP)* as an application protocol for distributed, collaborative client-server (also request- response) environment.

On other layers, a significant set of technologies are interconnected, many of them, in the development phase and specific areas of use, such as:

- *Resource Framework Description-Attributes (RDFa),* is an attribute level extension various XML or HTML based documents

- *Resource Framework Description- eXtensible Markup Language (XML RDF) - XML* is an expression syntax, which allows the transformation of the RDF graph as an XML document

- *Resource Framework Description Schema (RDFS)*, is a set of various classes enriched with enhanced knowledge-data models, which also provides basic elements for the description of ontologies, such as RDF vocabulary, intended for structuring RDF resources.

- *Web Ontology Language* (OWL) is a Semantic Web language designed to represent rich and complex knowledge about things, groups of things, and relations between things. OWL is a computational logic-based language such that knowledge expressed in OWL can be exploited by computer programs, e.g., to verify the consistency of that knowledge or to make implicit knowledge explicit. OWL documents, are known as *ontologies*. The goal of an ontology is to achieve common and shared knowledge that can be transferred between users and between the machines on which the applications are installed [8].

- *Query Language for RDF -SPARQL -* It is used to extract knowledge and information from ontologies, likewise the use of SQL to extract data from relational databases. SPARQL is used to express queries across diverse data sources, whether the data is stored natively as RDF or viewed as RDF via middleware. SPARQL also supports extensible value testing and constraining queries by source RDF graph. The results of SPARQL queries can be results sets or RDF graphs [11].

Beside semantic web stack technologies, which are constantly upgraded and expanded, in order to create, manipulate, and save the data for the needs of Semantic Web, it is advisable to use Semantic databases. These are predominantly NoSQL databases, such as MongoDB, which will be used in the following example of this paper.

2.2. SEMANTIC DATABASE MONGO DB COMPASS

Semantic databases as a concept are being developed and used to store semantic data models created using RDF and OWL. Semantic databases are NoSQL databases, which means non-relational database. In these database a specific way for storage and organization

of data is used, completely different from the relational databases' tabular relationships. In today's open market, e-commerce has developed its strategy to process various business applications. These applications are also used for customer e-commerce apps and require a scalable database for storing the data. NoSQL databases provides an efficient storage access and processing environment with horizontally scalable and replicable strategy over Relational Database Management System (RDBMS). Practically, it means that, unlike RDMBS, NoSQL are dominant and able to handle with: the massive use of online stores during seasonal discounts; the temporary nature of online e-shopping carts and online session information; storage and processing of completed orders over the internet [10]. Whenever there is a need for dynamic changes in the applied data model, in order to develop a successful e-business platform (e-commerce, some kind of e-store), it is recommended the use of document-based NoSQL databases. Mongo DB is an open-source scalable NoSQL document-oriented database, which is developed in C++. The working concept of "rows and columns" in relational databases is replaced with a model called "document". In MongoDB there is a way to combine the properties of the documents and the array of functionalities, which allows representing a hierarchical relationship at the level of a single record [11]. Further, through the functionalities of MongoDB Compass, the organization, creation, recharge and search of the database will be exposed.

3. AN EXAMPLE OF APPLYING MONGODB DATABASE TO THE NEEDS OF AN ONLINE STORE

The following example shows how to create, manipulate, upgrade and search through a semantic database, using MongoDB Compass. The required data models for the e-commerce process, in a fictitious online store, are created according to the rules of the semantic web. One of the possibilities for defining data models is to create an ontology, RDF, and triplets, previously in Protégé. According working prerequisites of online using, or installing Protégé, it is necessary to configure MongoDB. Actually, WebProtege uses MongoDB to store various configuration. MongoDB can be used in the future to store also other types of data. WebProtege will not start, if MongoDB is not installed. In the presented case, the data models, according to the semantic rules, are generated and defined when creating the database and the structures for each table, using only MongoDB Compass.

3.1. MONGODB COMPASS – INSTALATION AND CONFIGURATION

MongoDB Compass GUI, provides constant improvements and offers functionalities like, data visualization and exploration, data modification using power intuitive visual tools, debugging and optimization, understanding performance issues with visual explaining of plans, indexing decisions, and validation of documents. The Compass Plugin Framework is exposed as an API, making MongoDB Compass extensible by users [13]. Briefly described, the procedure for initial installation and configuration consists of the following steps:
- Installation of XAMPP, free open-source cross platform web-server solution stack, where server part is realized using Apache HTTP server. XAMPP is regularly updated to the latest releases of Apache, MariaDB, PHP and Perl. There are other technologies like database management system, MariaDB, which are replaced with MongoDB Compass in this working example.
- Installation of PHP, (7.2.1, is this case).

- Launching XAMPP and Apache Web server
- Download and installation of MongoDB Compas, Community server, and creating a folder for databases. Standard port for MongoDB is 21017. During this process, several basic PHP support files are created, which will be invoked at every MongoDB activation, some of them, with the ability to be customized.
- Launching MongoDB PHP Driver, for the interconnection needs between the two technologies, by making changes to the configuration files according to the given instructions in tutorials.

Some of these steps are repeated at every start of MongoDB and XAMPP server technologies, and they ensure consistency in operation, but also, if needed, supports the working changes.

3.2. AN EXAMPLE OF DATA MANIPULATION FOR E-COMMERCE NEEDS USING MONGODB COMPASS, XAMPP AND PHP

Through the following example, the intention is to explain the technological changes and perceptions of the new semantic web stack. The used semantic web technologies complement each other, rather than being replaced. After the installation and configuration, the working procedure takes place according to the steps described below.
- The XAMPP solution package is activated and Apache's web server is running, and from the XAMPP' main menu, the "Open application folder" is selected.
- Further, following the path   *htdocs/mongo*, there is access to several .php files that support MongoDB activities and complement them, such as:
    a. *db.php* - the configuration file for the local computer, as previously described, at port 27017 and the MongoDB driver.
    b. *index.php* - supports the creation of various searches and contains the information about created collections and documents (in this case, collection: ecommerce and defined and added document: products)
    c. *insert.php* - gives an overview of all entries (collection documents) and offers the function "Insert" to enrich existing collections with documents of the same or different format.
- The next step is MongoDB Compass' activating. The same options for searching for collections and options for inserting new documents and editing existing ones are available in MongoDB Compass. Any change in MongoDB reflects the changes in index.php, and insert.php., and in reverse.

By combining MongoDB, PHP, XAMPP, according Figure 2, the requirements for creating flexible documents have been solved, and data manipulation according to semantic rules has been made possible.
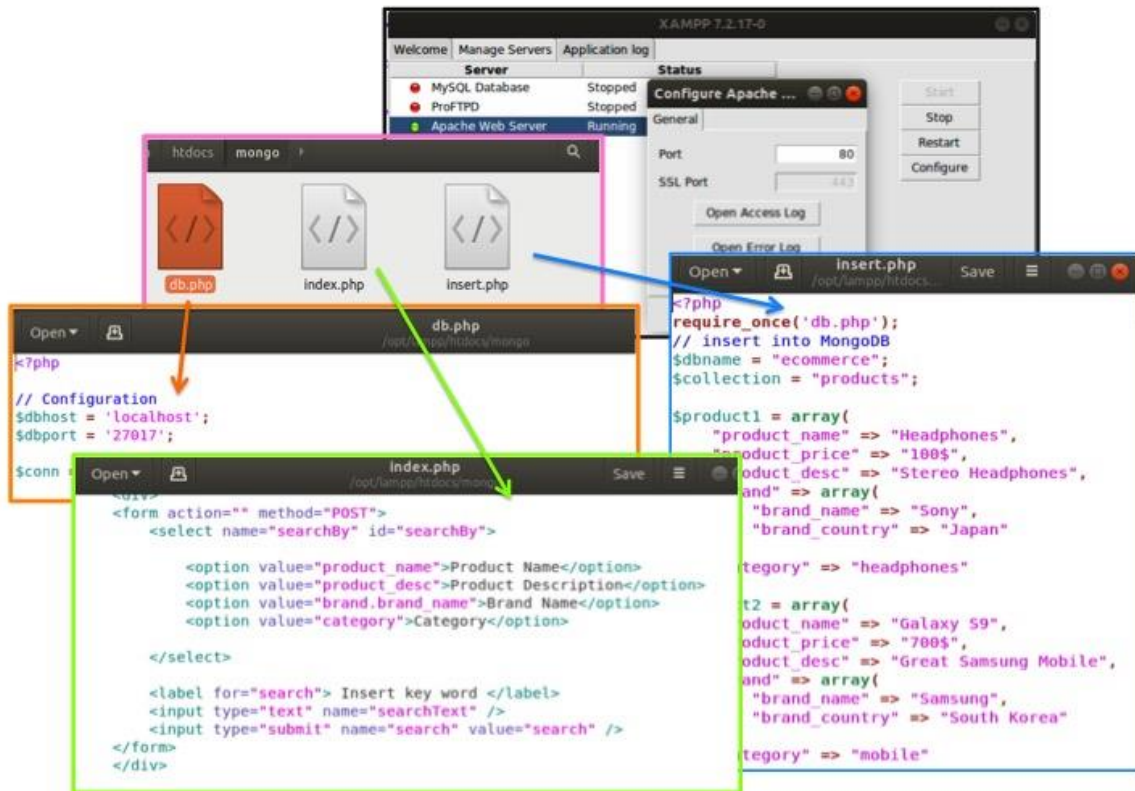
Figure 2. XAMPP, PHP, MongoDB technologies

As an illustrative example of the semantic technologies' use in e- commerce, a collection of products is created in MongoDB Compass. A data format is created called "ObjectID", which consist of the elements: "Product name", "Product price", "Product description", "Brand" (with two sub-elements "Brand name" and "Brand country") and "Category".
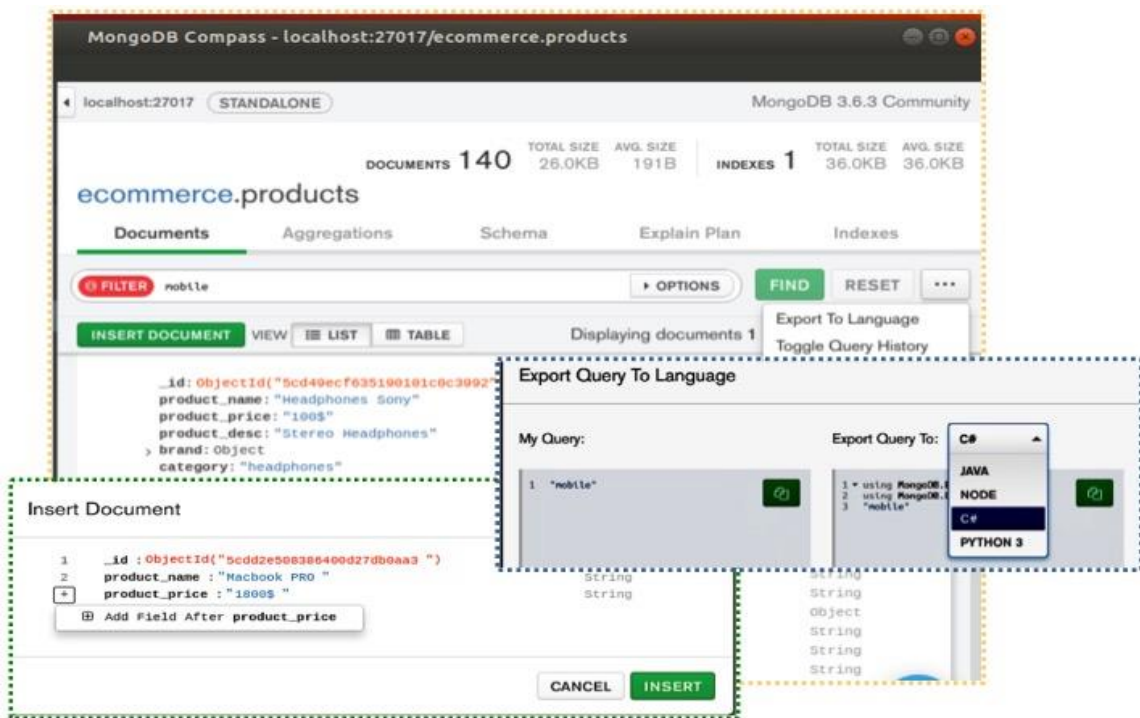
Figure 3. Basic functionalities of MongoDB Compass

The collection can be replenished with other documents of the same or different format, through the option "Insert Document". As needed, the elements of an existing document can be edited. There is also a possibility to define queries and export them to a few languages, such as C#, Java, Python 3, Node.js. Figure 3 shows the previously described functionalities in MongoDB Compass. Through the corrections and changes in the PHP code, the criteria for searching the database can be changed, as shown in Figure 4.
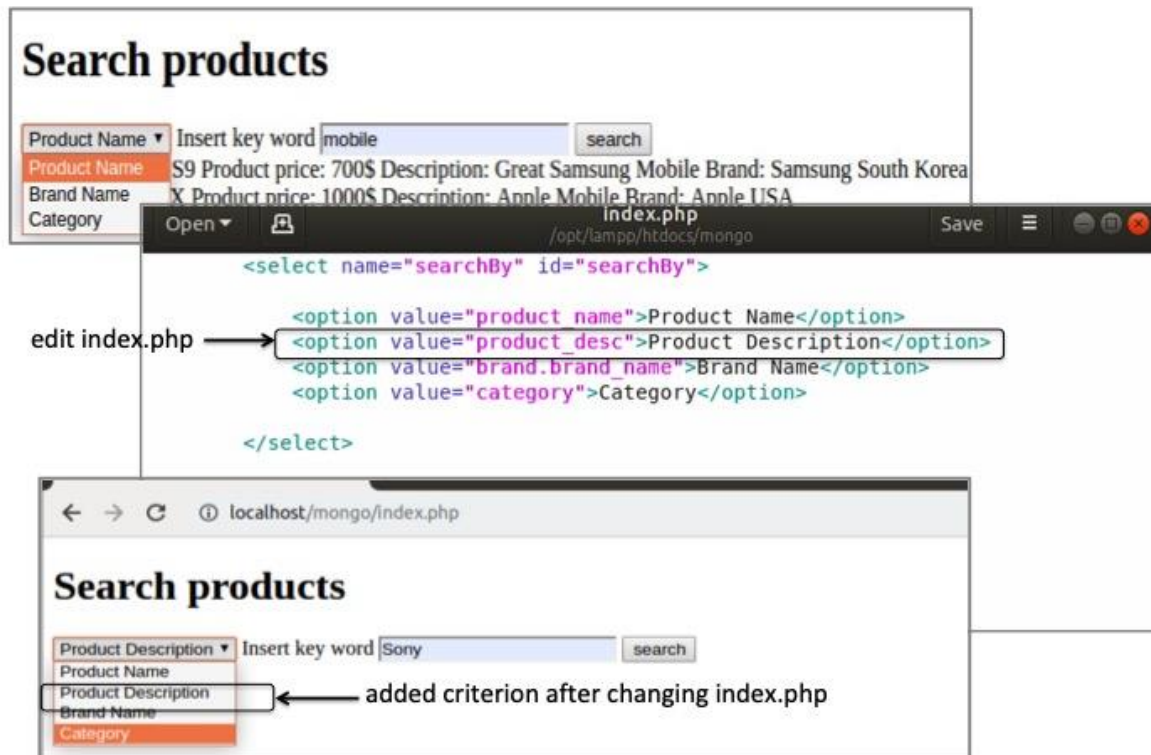
Figure 4. Querying MongoDB, through PHP

On the top left screenshot in the Search drop-down menu, can be searched for "Product name", "Brand name" and "Category". The central screenshot shows a deliberate modification of the PHP index.php file by adding the "Product Description" search option to the "Search" drop-down menu, according the previously described "ObjectID" document data format. On the last bottom-screen, the newly-added search item appears in the search menu, with the next execution of index.php in the browser.

Basically, MongoDB provides a JSON-based declarative query language consisting of two major components: the *find* query which retrieves documents matching a set of conditions, and the *aggregate* query based of processing pipelines, where each document of a collection passes through the stages of the pipeline, which leads to generating better-quality computations [13].

4. CONSCLUSION

Following the new restructuring of the global network, it comes to the conclusion that the technological basis for the wide realization of semantic web technologies in electronic commerce is created. Semantic technologies mark continuous growth, qualitatively and quantitatively, complementing and combining each other, in many different ways. Beside these features, they provide new ways of organizing and manipulating data that is relevant from a business point of view. In addition, there are several different ways of using and combining semantic web technologies, while designing and creating a specific solution. With the constant change and promotion of new technologies, the needs of e-business processes are followed, and existing solutions for creating, storing and searching data are being developed.

Previously reviewed and described technologies are just an instantaneous cross-section of the semantic web stack, and according that, the given example is one of the set of possible solutions. Therefore, the conclusion that is derived from the analyses made in this paper is that the semantic web technologies are best suited toward the informational needs of modern e-business processes, with an emphasis on e-commerce.

REFERENCES

1. Neha, V. et al. "E-commerce website ranking using semantic web minning and neural computing", International Conference on Advanced Computing Technologies and Applications (ICACTA-2015), 42-51.

2. Markoska R., Markoski A., "The Semantic Web Conceptual Solution and Working Ontology for Tourism Ecosystem", 8th International Conference on Information Society and Technology, ICIST, 2018, 327-332.

3. Zulqurnan A., "Semantic Web Mining in E-Commerce Websites", International Journal of Computer Applications , 137 (2),(201),123-131.

4. Heidari K., A Dehghan. "The impact of semantic web on e-commerc", World Academy of Science, Engineering and Technology 51, (2009), 17-31.

5. The World Wide Web Consortium (W3C) https://www.w3.org/standards/semanticweb/ , last visited 12.05.2019.

6. Semantic WEB technologies – Fundamentals, https://www.ontotext.com/ , last visited online 12.09.2019.

7. Polleres A., "Semantic WEB technologies: from theory to standards" 2010, available online at https://aic.ai.wu.ac.at/~polleres/publications/poll-2010aics.pdf

8. Taye M.M., "Understanding Semantic Web and Ontologies: Theory and Applications", Journal of Computing, 2 (6) , (2010) , 182-188.

9. Gandon F., "A Survey of the First 20 Years of Research on Semantic Web and Linked Data", Revue des Sciences et Technologies de l'Information - Série ISI: Ingénierie des Systèmes d'Information, Lavoisier, 2018.

10. Dharavat R., Ekaansh K., Shankar N.B., "Inclusion of e-Commerce Workflow with NoSQL DBMS: MongoDB Document Store", 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC).

11. Aboutorabi, S. H., at all "Performance evaluation of SQL and MongoDB databases for big e-commerce data", Computer Science and Software Engineering (CSSE), 2015 International Symposium, (1-7). IEEE.

12. Michel F., Faron-Zucker C., Montagnat J. (2016) A Mapping-Based Method to Query MongoDB Documents with SPARQL. In: Hartmann S., Ma H. (eds) Database and Expert Systems Applications. DEXA 2016. Lecture Notes in Computer Science, vol 9828. Springer, Cham.

13. MongoDB Compass, https://www.mongodb.com/products/compass , last visited online 13.05.2019.