INFORMATION
SYSTEMS IN
MANAGEMENT

Information Systems in Management (2016) Vol. 5 (3) 318–325

SEMANTIC TECHNOLOGIES BASED METHOD OF COLLECTION, PROCESSING AND SHARING INFORMATION ALONG FOOD CHAIN

DARIUSZ DOBROWOLSKI [a], ANDRZEJ MARCINIAK [b], GRZEGORZ BARTNIK [c],
PAWEŁ KAPŁAŃSKI [d], ZDZISŁAW ŁOJEWSKI [a]

[a] *Institute of Computer Science, Maria Curie Sklodowska University (UMCS)*
[b] *Faculty of Transport and Computer Science, University of Economics and Innovation in Lublin*
[c] *Faculty of Production Engineering, University of Life Sciences in Lublin*
[d] *Department of Applied Informatics in Management, Gdansk University of Technology*

In the paper the method of collecting, processing and sharing information along food chain is presented. Innovative features of that method result from advantages of data engineering based on semantic technologies. The source to build ontology are standards and regulations related to food production, and data collected in databases owned by food chain participants. It allows food chain information resources can be represented in semantic languages RDF/RDFS/OWL and form semantic database where data are easily integrated according to various criteria using automatic inference algorithms. Feasibility of proposed method is achievable with Ontorion Knowledge Framework.

Keywords: Food Chain, Semantic Technologies, Ontology, Enterprise Integration Platform, Ontorion, FluentEditor

1. Introduction

The concept of feasibility study defined by the European Union includes the operations carried out at the stage of formulation of the project, verification if the project has a good foundation for implementation and if it meets the needs of the targeted beneficiaries. The study should become a project plan. There must be identified and critically examined all the operational details of its implementation, such as commercial, technical, financial, economic, institutional, socio-cultural and environmental issues. The feasibility study allows to determine the financial and economic profitability, and as a result it creates a clear justification for the goal of the project.

Organizations belonging to the food chain form an ordered set from feed producers and primary products by manufacturers of foodstuffs, transport and storage operators and their subcontractors to the retail sales of nutritional products and services (including organizations indirectly related to the sector, such as manufacturers of machinery and equipment, packaging materials, cleaning ingredients and additives). The food chain also includes service organizations. According to ISO 22000: 2005 definition of food security refers to the presence of food-related risks at the moment of its consumption. Occurrence of risk referring to the food safety can appear at any stage of the food chain, so it is necessary to supervise appropriately the entire food chain.

Communication with the whole food chain is essential to ensure that all relevant risks relevant to the food safety are identified and adequately supervised at every stage. It means communication between organizations of the food chain towards its start and its end. Communication with suppliers and customers regarding identified risks and monitoring measures will help to determine precisely the requirements and expectations of the customer and the supplier (e.g. feasibility or the impact of certain expectations and requirements on the finished product).

The problems of Organizations is the lack of the documented information provided in an automatically processable format. Defining an uniform, commonly accepted framework for the proposed feasibility study, including methods of description of shared data and their availability and defining the behaviour of the system in a dynamic environment, will enable to solve a number of integration problems and enable to generate solutions based on an uniform standard. It will facilitate the availability and searching heterogeneous information resources and creation of interfaces. One of the purposes for which it is made an attempt to integrate data, is an improvement of processes in the Organizations from the food industry. As a result, it will partly ensure the required level of food safety, reduction of costs and increase of the efficiency of operations.

The diversity of the internal structure of information resources which are the subject of integration is important, but not the only problem faced by designers of integration solutions.

Among many issues, the following ones are in the centre of attention:

- variety of systems, query languages, integration strategies,
- redundancy, which makes the performance operations on data difficult, including aggregation,
- the problem of the quality and reliability of data.

There are many types of systems and kinds of relationships between organizations that require integration. Solutions in this area include shared systems, i.e. those that are owned by all cooperating organizations and are jointly managed, or in which one organization makes its resources available to other systems, loosely coupled systems, integrated systems, cooperating systems, and systems based on a common interface, such as a website.

2. Methodology

Method of collection, processing and providing information proposed by our Team regarding the risks for integrity of the food, existing in each segment of the food production chain takes into consideration the following assumptions [5, 6, 7, 8, 9, 10]:

- Risks arise in all segments of the food production chain. The structure of this chain is a trajectory in the network of biological, climatic, technological, economic and sociological interaction. Every interaction is connected with specific risks. The categorization of these risks (what are they and what are their property) is essential in risk management and food safety. Conceptualization of this complex system of risks should be formalized in a way that enables automated processing of information. This possibility is created by an ontological engineering of knowledge.
- Knowledge represented in the conceptual structures defined by ontology enables its automated processing by algorithmic inference processes. The stream of information along the food chain should be modular, i.e. local, separable and factorized. This means that in each segment of interaction networks, the specific information for a particular segment should be connected to the set of processing information. The entire chain is mapped in the information space, built according to the semantic technologies. The structure of semantic database is a modular graph, where each module represents a semantic subgraph, specific for a particular segment of the chain.
- To build ontology, the source of concepts and inter conceptual relations are standards and regulations: HACCP, BRC and IFS standards, ISO 9001, ISO

22000 - GHP, GMP, Codex Alimentarius, Council Directive 89/397 / EEC of 14.06. 89, the Council Directive No 93/43 / EEC of 14.06. 93, Regulation (EC) No 852/2004 of the European Parliament and the Council of 29 April 2004., Regulation (EC) No 853/2004 of the European Parliament and the Council of 29 April 2004 . We assume that every organization that is a participant in the food chain complies with the norms and standards and documents its activity.

- The system of information flow must be evolutionary, self-learning and self-organizing. The food production chain must be predictable and verifiable by independent, directed for this purpose system. Each part of integrated food must be connected with the identifying element directing to resources stored in the semantic database.

At present, there is no universal, uniform, widely-accepted model of data exchange. Multi-faceted diversity and wide range of issues for which you should find a solution when creating the integration system caused the creation of many proposals and systems that can be used in certain situations.

Feasibility study of the project is based on the available semantic technologies. Therefore the proposed method is a synthesis of relational technologies of databases that are commonly used by participants of the food chain and semantic technologies.

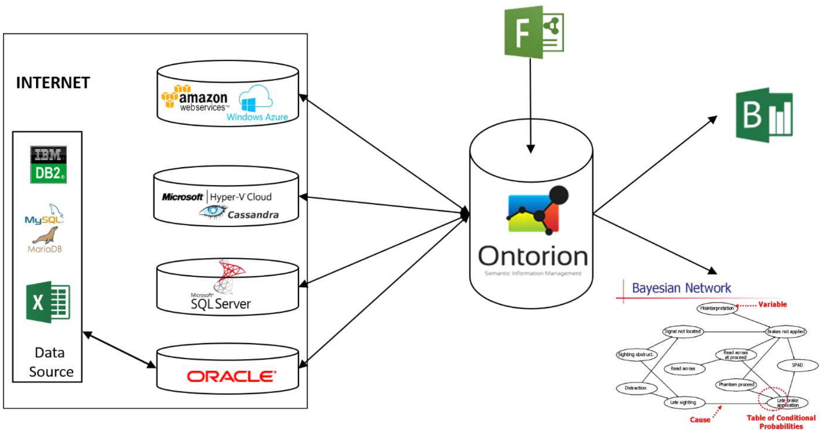

Figure 1. Flowchart of the project

The structure of such a system is shown in Fig. 1. Regardless of the structure of the data source, data is collected and processed to a uniform system of representation (RDF, RDFS, OWL), for example Oracle Spatial and Graph. Processed data is collected by ONTORION, the semantic server of knowledge. This server cooperates with Cloud systems.

2.1. Ontorion

Ontorion is a Distributed Knowledge Management System with Natural Language interfaces (CNL) [2] and a built-in rules engine. It is compatible with Web Ontology Language 2 (OWL2) [1] and Semantic Web Rule Language (SWRL) [4] and can be hosted in the Cloud or OnPremise environments. Ontorion is a family of products of server and client-side components for desktop and web allowing for the broad integration of custom software and existing corporate infrastructure. Ontorion performs real-time reasoning over the stream of data with the aid of an ontology that expresses the meaning of the given data (see Figure 2).

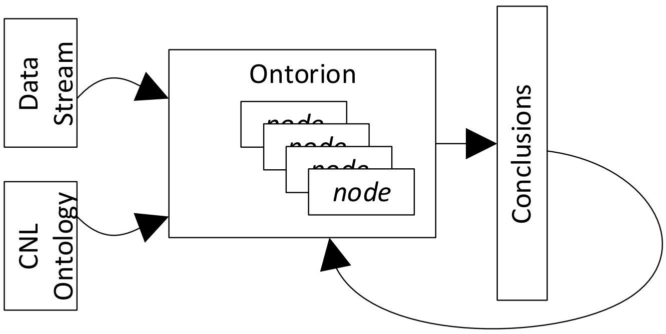

Figure 2. Ontorion - Knowledge Management System

Ontorion is a set of components equipped with algorithms that allows one to build large, scalable solutions for the Semantic Web. The scalability is realized by both the NoSQL, symmetric database and the ontology Modularization algorithm. Modularization algorithm splits the problem into smaller pieces that are able to be processed in parallel by the set of computational nodes, therefore; Ontorion is a symmetric cluster of servers, able to perform reasoning on large ontologies. Every single Ontorion Node is able to make the same operations on data. It tries to get the minimal suitable ontology module (component) and perform the desirable task on it. Symmetry of the Ontorion cluster provides the ability for it to run in the “Computing Cloud” environment, where the total number of nodes can change in time depending on the user needs.

2.2. FluentEditor 2014

FluentEditor 2014 [3], an ontology editor, is a comprehensive tool for editing and manipulating complex ontologies that uses CNL. FluentEditor, shown in Figure 2, provides a more suitable alternative for human users to eXtensible Markup Language (XML)-based OWL editors. Its main feature is the usage of Controlled English as the knowledge modelling language. Supported via Predictive Editor, it prohibits one from entering any sentence that is grammatically or morphologically incorrect and actively helps the user during sentence writing.

Controlled English is a subset of Standard English with restricted grammar and vocabulary in order to reduce the ambiguity and complexity inherent in full English.

Main features:

- CNL OWL implementation: The implementation of CNL OWL - FluentEditor grammar is compatible with OWL-DL and OWL2.
- OWL 2.0 full compliance: Full compliance with OWL 2.0 standard from W3C.
- OWL API: Compatible with OWL API, which allows it to be used in cooperation with other tools.
- SWRL compliance: The user can import existing ontologies from OWL files
- Dynamic referencing of external OWL ontologies: CNL documents can dynamically reference external OWLs from Web or disk.
- Predictive Edition Support: Users have enhanced support from the predictive editor.
- Built-in dictionary: The built-in dictionary makes it easier to avoid misspelling errors.

Among other features included, such as:

- Advanced user Interface, in order to open up semantic technologies for inexperienced users.
- In-place error resolving support - direct information about possible errors with hints on how to resolve them.
- Importing existing ontologies – users can directly import to CNL any external ontology.
- Ambiguity resolution - it keeps track of ambiguities of concepts and/or instance names imported from different external ontologies.

Data contained in ontologies can be much more diverse than data stored in databases. Extraction of information from the ontology is possible, either through direct query syntax, as in the case of a search, using Internet search engines, and determination (calculation) of the response by the rules of inference. One of the possibilities of creation of ontology of this field is the use of its representation in the form of a relational database. Ontological representation of information contained in relational schema requires answers to the questions: what exists in the present, represented by the relational schema, what is it and (classes, subclasses) and what are its properties (record fields). Ontology expressed by the OWL language is a repository of knowledge from which information can be extracted. Extracting information expressed in terms of ontology has characteristics both of extracting information stored in databases, as well as obtaining information from websites using web search engine.

From the point of view of modeling of complex manufacturing processes, where there are random factors, it is the ability to retrieve data from the graph database to algorithms of machine learning of Bayesian probabilistic networks. In order that, the production process of goods and services was methodologically and technologically integrated with information on production process and knowledge, it is needed to build a unified, interoperable information infrastructure for both processes. This means in practice the need to build into production process the advanced sensor fields and automated continuous acquisition of petabyte data streams. This requires the use of information technology known as Big Data, that is based on the collection of data in semantic networks.

3. Conclusion

Specialization and intensification of agricultural production causes the modern farms to become economic organizations of the increasing organizational and technological complexity. In the processes of changes in the management of agricultural production one can notice the phenomenon of organizational convergence, involving the tightening of the management methods and techniques used in agriculture and industrial production. The driving force behind this convergence is the development of information technology and the increasing importance of knowledge as a means of production. Modern production engineering increasingly uses the knowledge and information as a basic, elementary production resource. Currently extremely rapid development of semantic technologies can be seen, including the use of ontological knowledge engineering.

Semantic technologies of knowledge representation allow us to automate the integration process of knowledge from different sources. Currently, ontologies are the subject of research in a variety of research facilities, including engineering, information systems, knowledge engineering, natural language engineering, as well as in the theory of knowledge management.

Semantic technologies are necessary for the automated collection and processing of heterogeneous distributed intensive data streams. The main objective of the processing of such data is, in this case, their semantic integration and cognitive modeling, which results in machine-processable knowledge representation language. Semantic knowledge representation system allows for deep operationalization of pragmatics, understood as informational and therefore low energy impact on the course of production processes in the management and control of these processes.

REFERENCES

[1] P. Hitzler, M. Krotzsch, B. Parsia, P. F. Patel-Schneider and S. Rudolph, “OWL 2 Web Ontology Language Primer,” 20 październik 2015. [Online]. Available: http://www.w3.org/TR/owl2-primer/.

[2] Cognitum, “Ontorion Semantic Knowledge Management Framework,” Cognitum, 20 październik 2015. [Online]. Available: http://www.cognitum.eu/semantics/Ontorion/.

[3] Cognitum, “Fluent Editor 2014 - Ontology Editor,” Cognitum, 20 październik 2015. [Online]. Available: http://www.cognitum.eu/semantics/FluentEditor.

[4] W3C, “SWRL: A Semantic Web Rule Language Combining,” W3C, 20 październik 2015. [Online]. Available: http://www.w3.org/Submission/SWRL/.

[5] ISO, PN-EN ISO 22000: 2006 System Zarządzania Bezpieczeństwem Żywności - wymagania dla wszystkich uczestników łańcucha żywnościowego., 2006.

[6] ISO, PN-EN ISO 9000:2006 Systemy zarządzania jakością- Podstawy i terminologia., 2006.

[7] WE, Rozporządzenie (WE) Nr. 178/2002 Parlamentu Europejskiego i Rady z dnia 28 stycznia 2002 r. ustanawiające ogólne zasady i wymagania prawa żywnościowego, 2002.

[8] WE, Rozporządzenie (WE) Nr 852/ 2004 Parlamentu Europejskiego i Rady z dnia 29 kwietnia 2004 r. w sprawie higieny środków spożywczych, 2004.

[9] WE, Rozporządzenie (WE) Nr 853/ 2004 Parlamentu Europejskiego i Rady z dnia 29 kwietnia 2004 r. ustanawiającym szczególne przepisy dotyczące higieny, 2004.

[10] WE, Rozporządzenie (WE) Nr 854/ 2004 Parlamentu Europejskiego i Rady z dnia 29 kwietnia 2004 r. ustanawiającym szczególne przepisy dotyczące organizacji urzędowych kontroli, 2004.