IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Vol.-6, Issue-6, June 2018 E-ISSN: 2347-2693

A Comprehensive Review Of Web Semantic Technologies In Current World

Manas Kumar Yogi[1], G.Sai Sri Kavya[2]

[1]Department Of CSE, Pragati Engineering College, Surampalem, Dist:East Godavari,A.P., India
[2]Department Of CSE, Pragati Engineering College, Surampalem, Dist:East Godavari,A.P., India

Corresponding Author: manas.yogi@gmail.com, Tel.: +91 9966979279

Available online at: www.ijcseonline.org

Accepted: 03/Jun/2018, Published: 30/Jun/2018

Abstract— This paper diligently presents a comprehensive review of applications of web semantic technologies. Web semantics provides the backbone for information exchange among numerous diverse applications which are quite complex in themselves. The objective is to provide a controlled and scalable channel of knowledge transfer over the web .In this review paper we have limited our presentation to only three popular applications of web technology namely in digital libraries, legal domain, telecommunications. Our paper provides insights into few research projects currently taken up as well as challenges faced by researchers working in this field.

Keywords—Digital Libraries, Web Semantics, Ontology, eTOM, SID

I. INTRODUCTION

The comprehensive distribution of digital libraries over the last two decades is hardly remarkable. They offer remote approach to articles, journals and books with many users able to access the same document at the same time. With the use of search engines, they make it feasible to locate specific information more rapidly than ever is possible in physical libraries. Scholars, and others, are able to access rare and precious documents with no threat of damage. However, challenges persist if all the benefits are to be visualized. Interoperability between distinct libraries, or even between distinct collections in the same library, is a problem. At the semantic level, distinct schemas are used by distinct library databases. Search and retrieval need to be made simple, in part by offering each user a consolidated view of the naming of digital objects across libraries. User interfaces need to be enhanced, in particular to face the challenge of large information collections. This paper elucidates the state-of-the art in digital library research, and in particular the application of semantic technology to encounter the challenges imposed. St Succeeding sections go into detail, but it is clear that the challenges defined above regulate closely with the intention of semantic knowledge technology. The ontology mediation techniques are specially motivated by the challenge of interoperability between heterogeneous data sets, and of providing a standard view to those data sets. Semantic information approach offers enhanced ways to search for and browse information and, over an understanding of the correlation between documents, to enhance the user interface. Semantic access to information depends in turn on the supporting technologies described in the preceding papers; while the creation and maintenance of ontologies in digital libraries create problems of ontology management which require new insights into ontology engineering. The analysis is illustrated with a particular case study in which semantic knowledge technology is being introduced into the BT digital library. This provides an opportunity not just to trial the feasibility of the technology, but also to gauge the users' reactions and better understand their requirements. Finally, it should be remembered that digital libraries are themselves a particular form of content management application. Much of what is being learned here is relevant in the wider context of intelligent content management.

II. DIGITAL LIBRARIES: THE STATE-OF-THE-ART

Several working digital libraries are academic and make information easily available. Some examples are given in the section below describing digital library research. Others are commercial, such as the ACM digital library (http://portal.acm.org/dl.cfm), which contains material from ACM journals, newsletters and conference proceedings. Others, such as BT's digital library which we describe below, are for use within particular organizations. Another category of digital library exists for the explicit purpose of making material freely available. A well-known example of this is Project Gutenberg (http://www.gutenberg.net) which, at the time of writing in autumn 2005, has around 16 000 'eBooks' and claims to be the oldest producer of free e-books on the Internet. Correspondingly, the Open Library web site (http://www.openlibrary.org/ toc.html) created by the Internet Archive, in conjunction with organizations such as Yahoo and HP, to 'determine how books can be represented on-line' and 'create free web approach to important book collections from around the world'. A great deal of digital library software is openly feasible. One of the best known projects is the Greenstone digital library (http://www. greenstone.org) which is available in a wide variety of languages, Greenstone is sustained by UNESCO and, amongst other applications, is used to circulate practical information in the developing world. Another example is OpenDLib (http://opendlib.iei.pi.cnr.it), which is designed to support a distributed digital library, with services all over the Internet. A recent development from Google sees the world of the public domain search engine encroaching that of the digital library. Google Scholar (http://scholar.google.com/) provides access to 'peer-reviewed papers, theses, books, abstracts and other scholarly literature'. It uses the same technology as Google uses to access the public Web and applies this to on-line libraries. This includes using Google's ranking technology in order to search results by purpose. In a similar initiative, Yahoo is working with publishers to provide access to digital libraries. Considered 1995 a workshop held inferior to the auspices of the U.S. Government's Information Framework Technology and Applications

A. Research challenges identified for digital libraries:

1. Interoperability: At specific level this is about the interoperability of software and systems. At a deeper level, however, it is about semantic interoperability through the mapping of ontologies. Indeed 'deep semantic interoperability' has been identified as the 'Grand Challenge of Digital Libraries'

2. Description of objects and repositories: This is the requisite to organize common schema to enable distributed search and retrieval from distinct sources. Effectively, how can we create an ontology for searching and browsing into which we can map individual library ontologies? Going further, how can we enable individual users to search and browse within the context of their own personal ontologies?

3. The collection and management of nontextual information: This involves issues relating to the management, collection and presentation of digital content beyond multiple generations of hardware and software technologies. Moreover, libraries are more than collections of words, but are increasingly rich in audiovisual material, and thus raises new research challenges.

4. User interfaces: We need best ways to navigate large information collections. One approach is through the use of visualization techniques. The use of ontologies not only helps navigation but also provides a basis for information display.

5. Economic, Social and Legal Issues: These include digital rights management and 'the social context of digital documents'.

Semantic technology also impacts , for example through improving knowledge sharing in social groups. The need for interoperability across heterogeneous data sources is voiced by many authors. A recent U.S. workshop on research directions in digital libraries analyzed a number of basic themes for long-term research, of which one is interoperability, which describes as 'the grail of digital libraries research since the early 1990s'. A number of the other themes reiterate the need to overcome heterogeneity.

As implied by Lynch and Garcia-Molina (1995), the topic of digital libraries has attracted significant research activity since the 1990s. Some work has been with very precise goals. For example, the Alexandria Digital Earth Project (http://www.alexandria.ucsb.edu/), at the University of California, is concerned with geospatial data, whilst other projects have investigated areas such as medical information, music and mathematics. Another Framework project, Sculpteur 5 (http://www.sculpteur- web.org), used semantic technology for multimedia information management. The target domain is that of museums. An ontology, with associated tools, has been created to describe the objects, whilst a web crawler searches the Web for missing information. Currently, the Framework Programme is sponsoring significant research in the area of digital libraries. which 'Knowledge Extraction and Semantic Interoperability' is one.

B. Implementing Semantic Technology in a Digital Library

A well-designed ontology is a prerequisite for a notable semantic operation. Within SEKT we are affiliating a layered advent. In the lower layers we have a general ontology, which we call Proton (PROTo Ontology, http://proton.semanticweb.org). The classes in this ontology are a combination of very general, for example Person, Role, Topic, Time- Interval and classes which are more special to the world of business, for example Company, PublicCompany, MediaCompany.

A strength of an approach based on the use of an ontology language such as OWL, is the ability to receive distributed ontology creation activities, for example through defining equivalences. Nonetheless, where possible the creation of duplicate ontological classes should be deflected and where applicable we make use of prevailing well-established ontologies, Within PROTON there is a class, 'Topic'. Each topic is an instance of this class. However, usually a topic will be a sub-topic of another topic, for example in the sense that a document 'about' the former should also be regarded as being about the latter. Since topics are instances, not classes, we cannot use the inbuilt subclass property, but must define a new property subTopic. Such a relationship must be defined to be transitive, in the sense that if A is a sub-topic of B and B is a sub-topic of C, then A is also a sub-topic of C. This approach, based on defining topics as instances and using a subTopic property rather than defining topics as classes and using the sub-class relation, is chosen to avoid problems in computational tract- ability. In particular, this facilitates us to stay within OWL DL.

III. ONTOLOGIES FOR LEGAL KNOWLEDGE

Legal ontologies are different from other domain ontologies in two ways. On one hand, although legal statutes, legal judgments, or jurisprudence are written both in natural and technical language, all the common sense notions and connections among them, which people use in their everyday life, are symbolized in the legal domain. On the other hand, the strategy of ontology building must take into account the peculiar model of law that has been chosen. This occurs in a middle-out level that it is possible to skip in other ontologies based in a more contextual or physical environment. Therefore, the modelling process in the legal field usually requires an intermediate level in which several concepts are implicitly or explicitly related to a set of decisions about the kind of law, the nature of language used to personify legal knowledge, and the specific legal structure dealt by the ontology. There is an interpretative level that is frequently linked to general theories of law. This intermediate level is a well- known layer between the upper top and the domain-specific ontologies, especially in 'practical ontologies.' We may also inherently find this perception between an ontology layer and an application layer in cognitive modelling, in which categories, concepts and instances are distinguished. But the most striking feature of the legal ontologies constructed so far is that the intermediate layer is explicitly occupied by a kind of high conceptual constructs provided by general theories of law instead of empirical or cognitive findings.

A. Legal Ontologies: State of the Art

At present, many legal ontologies have been built. One prevalent way of describing the actual state of the art is to identify the main current legal ontologies :

- LLD [Language for Legal Discourse: based on atomic formula, rules, and modalities;
- NOR :based on agents behavioural invariants and realizations;
- LFU [Functional Ontology for Law:] based on normative knowledge, world knowledge, responsibility knowledge, reactive knowledge, and creative knowledge;
- FBO [Frame-Based Ontology of Law], positioned on norms, acts, and confession of concepts;
- LRI-Core Legal Ontology;positioned on objects, processes, physical entities, mental entities, assistants, and communicable acts;
- IKF-IF-LEX Ontology for Norm Comparison based on agents, institutive norms, instrumental provisions, regulative norms, open-textured legal notion, and norm dynamics.

At the moment, thirteen distinct legal ontologies have been identified, correlating to 10 years of research.

The legal ontologies characterized above have been built up with many purposes: information retrieval, statute retrieval, normative linking, knowledge management, or legal reasoning. Although the legal domain remains very sensitive to the features of peculiar statutes and regulations, some of the Legal-Core Ontologies (LCO) are meant to share a common kernel of legal notions. LCO remain in the domain of a general knowledge shared by legal theorists, national, or international jurists and comparative lawyers. However, our data indicate that there is a kind of specific legal knowledge, which belongs properly to the legal and judicial culture, and that is not being captured by the current LCO.

Professional knowledge comprises propositional knowledge (knowing that), procedural knowledge (knowing how), personal knowledge (intuitive, pre-propositional), and principles related to morals or deontological codes. Judges, prosecutors, and other court staff share only a portion of the legal knowledge (mostly, the legal language and the general knowledge of statutes and previous judgments). But there is a portion of this legal knowledge, the knowledge related to personal behaviour, practical rules, corporate beliefs, effect reckoning, and perspective on similar cases, that persists implicit and tacit within the relation among judges, prosecutors, attorneys, and lawyers. Technically speaking, these issues are not complicated. However, they are troublesome to solve. The judges' original question cannot be solved by simply pointing out a peculiar statute or legal doctrine. This is not only an problem of normative information retrieval. What is at stake here is a distinct kind of legal knowledge, a professional legal knowledge. In this regard, the design of legal ontologies requires not only to represent the legal, normative language of written documents (decisions, judgments, rulings, partitions), but also the professional knowledge categorized out from the daily practice at courts. From this point of view, skilful of a legal subject (such as e.g., gender violence) affiliates a peculiar knowledge of: (i) statutes, codes, and legal rules; (ii) professional training; (iii) legal procedures; (iv) public policies; (v) everyday routinely cases; (vi) practical situations; (vii) people's most common reactions to previous decisions on similar subjects. This Professional Legal Knowledge (PLK) is: (i) shared among members of a professional group (e.g., judges, attorneys, prosecutors...); (ii) learned and conveyed formally or most often informally in specific settings (e.g., the Judicial School, professional associations—the Bar, the Judiciary, etc.); (iii) expressible through a mixture of natural and technical language (legalese, legal slang); (iv) non equally distributed among the professional group; (v) non homogeneous (elaborated on individual bases); (vi) universally comprehensible by the members of the profession (there is a sort of implicit identification principle). Professional knowledge is a context-sensitive knowledge, tied up in courses of action or practical ways of behaving. In this sense, it implies: (i) the ability to discriminate among related but different situations; (ii) the practical attitude or disposition to rule, judge, or make a decision; (iii) the ability to relate new and past experiences of cases; (iv) the ability to share and discuss these experiences with the peer group.

3.2. Ontology of Professional Judicial Knowledge (OPJK)

The OPJK is studied from the competency questions imposed by the judges in their interviews. Modelling this professional judicial knowledge required the confession of this knowledge, as it was recognized by the judge. The OPJK has, currently, 700 terms, mostly relations and instances as a result of a choice to reduce the concepts at the class level when possible. Some top classes of the domain ontology identified are: CalificacionJurdica [LegalType], Jurisdiccion [Jurisdiction], Sancion [Sanction], Acto [Act], (which includes as subclasses ActoJurdico (LegalAct), Fase [Phase], and Proceso [Process]). These latter classes contain those taxonomies and relations related to the different types of judicial procedures (both, criminal and civil, or private) and the different stages that these procedures may have (period of proof, conclusions, appeal, etc.). The introduction of the concept Rol [Role] allowed for the specification of different situations where the same agent could play distinct parts. In the case of OPJK, the class Rol contains the concepts and instances of procedural roles [RolProcesal] that an agent might play during a given judicial procedure.

IV. A SEMANTIC SERVICE-ORIENTED ARCHITECTURE FOR THE TELECOMMUNICATIONS INDUSTRY

This section will describe the advantages of semantically described web services in the context of an SOA. In order to do this, the limitations of current web services are primarily considered. Web Services are broadly described using XML-based standards namely WSDL (which allows one to describe a Web Service in terms of what it does and what its inputs and outputs are), UDDI (which is a centralized registry allowing one to discover Web Services) and SOAP (which is a protocol allowing one to execute services). In addition to these low-level standards, work is in progress to build standards that access services to be combined into a workflow, for example WS-BPEL (Web Services — Business Process Execution Language) (IBM, 2005) and also to define acceptable message exchange patterns and contents, for example ebXML. However, no bit of these standards provide a way to describe a Web Service in terms of explicit semantics. For a given service you might want to describe: What kind of service it is; What inputs it requires; What outputs it provides; What needs to be accurate for the service to accomplished (pre-conditions); What becomes accurate once the service has carried out (post-conditions); What effect the service has on the state of the world (and/or the data it consumes and provides).

The fundamental of these prerequisites is partly addressed by UDDI in that a category and human readable description can be referred to a web service in a registry to aid discovery. This provides only less support for automated discovery since a computer do not understand the description or what the category means. The second and third of these prerequisites are partly addressed by WSDL in that XML tags can be attributed to inputs and outputs. A computer can easily match these but again has no concept of their meaning or relationship to other pieces of data. Typically, most of the hard work is left to the human user who must interpret the descriptions provided to the best of his or her abilities. Services can be elucidated semantically by relating them to ontologies. Ontologies provide a common view of a domain that can be interpreted by machines. Thus ontologies can describe kinds of services, the data they consume and accommodate, the processes that services are part of and, equally importantly, the relationships between all of the above. The explicit relations between services and ontologies is the key element for Semantic Web Services. It is envisaged that this will enable:

- Improved service discovery: Semantic Web search technology confess users to explore on ontological concepts rather than by keywords. A simple keyword search only finds where a peculiar term occurs, and does not give particulars about its context or relationship to other information. Ontological searches utilize the structured way that information is modelled to allow more powerful searches, such as the ability to query attributes or relationships between concepts.
- Re-use of service interfaces in different products/settings: Services that are described semantically can easily be discovered, understood and applied thus minimizing the need to create new services that serve the similar purpose.
- Simpler change management: Changes to models and services are inevitable over time. The key thing is to minimise the knock-on effect of change or at least manage it. A semantic approach will significantly reduce the overhead and simplify the process.
- A browse able, searchable knowledge base for developers (and others): In aggregation with the example given above for simpler change management, semantically described services and ontologies enable a knowledge base to be constructed.

- Semi-automatic service composition: Given a high level goal in which we wish a service or set of services to accomplish, expressed in terms of an ontology, it is possible to carry out decomposition into component parts and then match these components with appropriate services.
- Mediation between the data and process requirements of component services: Often there is need for two or more services to interact even though their communication requirements are semantically the same but syntactically different (they may require different message exchange patterns or different data formats). In this case it should be possible to automatically construct a translation between message data elements that allows the services to communicate.

A. SEMANTIC MEDIATION

The role of mediation in supporting an SOA has been noted. Mediation is typically accomplished through the use of mediators, that is components which enable heterogeneous systems to interact. In a practical sense, mediators have been realized as pieces of program code that perform point-to-point, low-level translations. Although such mediators satisfy the short-term goal in that they allow two systems to talk to each other, they suffer from maintainability and scalability problems. In general, it is not likely to be viable to automate their application in a dynamic environment because of their close coupling with the implementation. Semantic Mediation enables a more dynamic approach through the use of ontologies, which provide consensual and formal conceptualization of a given domain. 'Mediators can be used to convert from a source implementation interface to that of a target implementation. Modelling the processes and data in the source and target interfaces using ontologies, enables the definition of relationships between semantically equivalent concepts. The mediator can use these relationships to dynamically map between the source and target'.

V. ONTOLOGIES IN TELECOMMUNICATIONS

The Telecommunications Industry is probing means to strengthen interoperability among the many systems needed to run and manage a telecommunications network. One such approach is the New Generation Operations Systems and Software (NGOSS) initiative from the TeleManagement Forum (TeleManagement Forum, 2005a). NGOSS is an combined framework of industry agreed specifications and guidelines which include a common information and data model for systems analysis and design, and a process framework for business process analysis. NGOSS is intended to allow easier integration of the Operational Support Systems (OSS) software used to provision, bill and manage network-based products and services. Part of the work of NGOSS is to produce standards for Next Generation Networks (NGNs). Presently telecommunications companies have many distinct networks for distinct services (e.g. PSTN, Leased Line) that require managing and maintaining individually. This requires hundreds or even thousands of different bespoke system for each network to permit billing, maintenance, trouble reporting etc. Telco's are moving towards a centralized IP-based core to their networks, where many network services can be provided over one core network. This should lead to considerable cost savings and greatly enhance flexibility and efficiency in providing network services. NGOSS has identified that the use of SOA will be important in managing the NGNs as the benefits offered by SOAs fit well into the dynamic and highly flexible architecture that NGNs offer. The critical features of an SOA are captured in the NGOSS principles:

- Shared Information Data Model: NGOSS components implement and use a defined part of the Shared Information/Data Model (SID).
- Common Communications Vehicle: Reliable distributed communications infrastructure, for example software bus integrating NGOSS components and workflow.
- External Process Control: Separation of End-to-End Business Process Workflow from NGOSS Component functionality.
- Business Aware NGOSS Components: Component services/functionality are defined by NGOSS Contracts.

The work of the TeleManagement Forum in evolving a framework for Next Generation OSS can be seen as ontology building in that NGOSS gives a level of common approach for a particular domain of interest. NGOSS is applicable as a toolkit of industry-agreed specifications and guidelines that cover key business and technical areas including Business Process Automation and Systems Analysis and Design. The former is delivered in the enhanced Telecom Operations Map (eTOM TM) and the latter is delivered in the SID. The eTOM maintains a framework that allows processes to be assigned to it. It describes all the enterprise processes needed by a service provider. The SID provides a accepted vocabulary allowing these processes to communicate. It identifies the entities involved in OSS and the relationships between them. The SID can therefore be used to identify and describe the data that is consumed and produced by the processes.

A. eTOM

The eTOM can be scrutinized as a Business Process Framework, since its goal is to classify the business activities embodied in process elements so that these elements can then be united in many distinct ways, to implement end-to-end business processes (e.g., billing) which deliver value for the customer and the service provider. The eTOM can be dissolved to lower level process categories, for example 'Customer Relationship

Management' is decomposed into a number of categories, one of which is 'Problem Handling'. This is then decomposed further into categories such as 'Track and Manage Problem'. It is to these lower level categories that business specific processes can be mapped. eTOM uses hierarchical disintegration to structure the business processes. Process elements are formalized by means of a name, a description, inputs/outputs and a set of known process linkages (i.e., links to other relevant categories). The eTOM supports two different outlooks on the organization of the detailed process elements:

1. Horizontal process groupings, in which process elements depict functionality that spans horizontally across an enterprise's internal organisations (e.g., market, product, customer and service management etc.).

2.Vertical process groupings, in which process elements are grouped within End-To-End processes (e.g., fulfilment, assurance etc.) adepted by the Service Provider enterprise. The eTOM Business Process Framework is defined as typically as possible, so that it is separated of organization, technology and service.

B. SID

The SID is much more complex than the eTOM in both its aims and form. It gives a data model for a number of domains depicted by a aggregation of concepts known as Aggregate Business Entities. These use the eTOM as a focus to regulate the pertinent information to be modelled. The SID models entities and the relations between them. For instance a 'customer' is defined as a subclass of 'role'. It contains attributes such as 'id' and 'name'. It is linked to other entities such as 'CustomerAccount' with an association 'customerPossesses'.

VI. CONCLUSION

. The future Semantic Web will include a wide variety of heterogeneous resources. A Semantic Grid which effectively subsumes the Semantic Web and includes resources ranging from powerful computational resources to sensor networks. Amongst these will be the components of a digital library. Yet the digital library as an identifiable entity may have ceased to exist. Instead the user of the Web will see a network of resources, of varying provenance, trustworthiness and cost. Much will be free, but where payment is justifiable, then it will be required. The introduction of information technology should always be accompanied by a redesign of business processes. Our technology must be seamlessly integrated into the systems which support a user's work; and we must seek to go beyond the limitations of our paper-based metaphors and truly exploit the power of the technology. To achieve all this, significant research is still needed. Just as in other papers' authors have stressed the need for more research into the core semantic technologies, so here we stress the need for more research into exploiting those technologies to create the digital libraries of the future. Semantic Web has proved to be very useful in improving the knowledge management skills of the recently appointed judges. The legal applications of web technologies are designed not only to be accurate and technologically advanced, but also to fulfill the specific requirements of professional judges. It is designed to be efficient, extensible, customizable, and scalable. It makes use of incremental search as a process of narrowing the solicited FAQ set. It uses a variety of pluggable searching algorithms.

REFERENCES

[1]K. Aberer, P. Cudre-Mauroux, and M. Hauswirth. The Chatty Web: Emergent Semantics Through Gossiping. In Proceedings of the 12th International World Wide Web Conference (WWW'03), Budapest, Hungary, 20-24 May 2003. ACM.

[2] S. Alexaki, V. Christophides, G. Karvounarakis, D. Plexousakis, K. Tolle, B. Amann, I. Fundulaki, M. Scholl, and A.-M. Vercoustre. Managing RDF Metadata for Community Webs. In Proceedings of the ER'00 2nd International Workshop on the World Wide Web and Conceptual Modeling (WCM'00), pages 140–151, Salt Lake City, Utah, 9-12 October 2000.

[3] Yigal Arens, Chun-Nan Hsu, and Craig A. Knoblock. Query Processing in the SIMS Information Mediator. In Advanced Planning Technology, California, USA, 1996. AAAI Press.

[4] Ricardo Baeza-Yates and Berthier Ribeiro-Neto. Modern Information Retrieval. Addison Wesley Longman, 1999.

[5] R. Beckwith and G. A. Miller. Implementing a Lexical Network. Technical Report 43, Princeton University, 1992.

Authors Profile

Mr. Manas Kumar Yogi pursued Bachelor of Technology from VR Siddhartha Engineering College,Vijayawada,A.P. in 2006 and Master of Technology From Malla Reddy College Of Engineering And Technology in year 2012. He is currently working as Assistant Professor in Department of Computer Science Engineering , Pragati Engineering College (Autonomous), Surampalem, East Godavari District, since 2014. He is a member of IEEE & ACM since 2014. He has published more than 50 review,research papers in reputed international journals ,conferences including IETE sponsered conferences. His main research work focuses on Software Engineering, Distributed Computing, Cloud Security and Privacy, Big Data Analytics, , IoT and Computational Intelligence based optimisations. He has 8 years of teaching experience and 2 years of software industry Experience.

Miss Godaverty Sai Sri Kavya pursued her Intermediate at Sri Chaitanya Junior College Kakinada in 2015. She is presently pursuing her final year B.Tech. in Computer Science Engineering at Pragati Engineering College Surampalem ,East Godavari District,A.P.. Her areas of interest include IoT and Web Technologies .

