Role of Semantically Enriched Ontologies for Information Retrieval in Indian E-Governance

Namrata Rastogi[1], Dr. Parul Verma[2]
Dept. of Information Technology
Amity University, Lucknow, India

***Abstract*—A major driving force required for Electronic Government (E-Government) is the dominant concept of Semantic Ontologies. Semantically enriched Web Ontologies aims to provide a superlative level of automation and better performance of government functions and services.**
Due to the heterogeneous structure of the government sector, achieving the interoperability and integration of government departments is the biggest challenge for a perfect e-Government. Therefore, the amalgamation of E-Government and Semantic Ontologies is the need of the time.

In this paper, we present a role of semantically enriched ontologies in Indian E-governance. The aim of our approach is to improve the E-government- citizen relationship by providing automated information retrieval to citizens of India and reducing the heterogeneity of various government divisions and thereby optimizing E-Government functions and services.

KEYWORDS - semantic web, ontology, E-government, information retrieval

I. INTRODUCTION

Today, we are living in a world where Information Communication Technology (ICT) is becoming a part of almost every spheres of human activity at an unprecedented rate. In the past few years, ICT is being used for providing information related to public services, improving managerial effectiveness and promoting democracy, which is commonly termed as e-government.
E-Government in India aims to develop computer based solutions to support the interaction of citizens with the public Sector and improve their participation in the public and social life [3].

Semantic technologies aim to provide advanced E-Government solutions and hence require development of E-Government based ontologies.

One of the key objectives of semantic technologies applied to E-Government is to provide better services while enhancing the participation of citizens to the public life. Semantic technologies, to be effective, need to be based on sound ontologies that need to be shared within the participating social communities [5].

A. The issues in manual government:

- Long waiting queues in government offices – The total time a citizen spends getting his work done in a satisfactory manner including waiting time in queues, total response time given by the government official on duty.
- Corruption and quality of service – The total bribe paid at various desks to get the same work completed in an orderly and timely manner and hence degrading average quality and trust of service between citizen and government officials.
- Redundancy of similar data of citizens' in various government departments – Recoding and maintaining similar data records of each citizen by every service departments like Electricity department, LPG connection, Income tax department, EPIC card, etc. and thus creating redundancy of similar data records.
- Faulty and poor overall governance – No quality maintained in establishing transparency, quality of feedback, fairness of treatment, accountability levels and reducing corruption.
- Lack of Interoperability, heterogeneity of government sectors

II. SEMANTIC WEB ONTOLOGIES

Tim Berners-Lee, who invented the WWW and has worked on the Semantic Web in 1980s, states that the Semantic web "is not a separate Web but an extension of the current one, in which information is given a well-defined meaning, better enabling computers and people to work in cooperation." [2]. Thus, the Semantic Web [16, 19] is distinguished by a more meaningful representation of information for humans and computers, providing a depiction of its contents and services in machine-readable format; it also enables services to be automatically annotated, discovered, published, advertised and composed. It thereby facilitates interoperability and the sharing of knowledge over the Web.

The Semantic Web is distributed and heterogeneous by nature an hence has two visions of its future development, the first is to improve its usability by collaborating various distributed

working departments and the second to make Web, machine friendly and more automated in its working and understanding. Its main goal is therefore to make information on the Web accessible and understandable by humans and computers.

Semantic Web [1] is a group of methods and technologies that allow machines to understand the meaning (semantics) of the information which is residing on World Wide Web. Furthermore, Semantic Web is providing the share and reuse of knowledgebase representation, global linked representation of knowledgebase [2] to stipulating semantic based access of web resources and extracting information from these resources.

A. Benefits of Semantic Web

- Knowledge sharing
- Automation of Web
- Machine friendly
- Reduction of data redundancy
- Reusability of web content
- Searching targeted information using semantics is extremely easy
- Quick and efficient information retrieval based on semantics and ontologies.

Semantic Web is implemented using *Ontology* which is a formal explicit description of concepts in a given domain, properties of each concept describing various features and attributes of the concept, and restrictions or relationship among classes or concepts. Ontology along with a set of individual instances of classes forms a knowledge base.

In the field of the Artificial Intelligence, Neches [6] was the first to define ontology, and he did it as follows: "Ontology defines the basic terms and the relations that include the vocabulary of a specific area, in addition to the rules to combine terms and relations to define extensions to the vocabulary".

Studer and colleagues [8] explained Ontology as follows: "Ontology is a formal, explicit specification of a shared conceptualization. Conceptualization refers to an abstract model of some phenomenon. Explicit means that the type of concepts used, and the constraints on their use are explicitly defined. Formal refers to the fact that the ontology should be machine-readable. Shared reflects the notion that an ontology captures consensual knowledge, that is, it is not private of some individual, but accepted by a group".

The goal of ontology is to achieve a common and shared knowledge that can be transmitted between people and between application systems. Thus, ontologies [7] play an important role in achieving interoperability across organizations and on the Semantic Web [16], because they aim to capture domain knowledge and their role is to create semantics explicitly in a universal manner, giving the basis for agreement within a domain.

III. RELATED WORK

Development of various E-Governance applications using Semantic Web technology is in early stages. Researchers all over the world are into development of semantic based e-Governance. But it is not an easy task to integrate the Semantic Web technology with the much complexed E-Governance domain. Though there are few researchers which have successfully completed their projects, but still there is a long way to go in executing the prototypes.

This section will discuss about few such research projects going across the world which talk about problems like interoperability, heterogeneity in e-government sector and propose the solution of semantic web and Ontology to get rid of above mentioned issues and increase the usability of the solution.

The major problem found while integrating information systems in E-government is that of interoperability. Mustafa Jarrar, Anton Deik, Bilal Farraj from Palestine [9] in their work presented the case of the Palestinian Interoperability Framework 'Zinnar', which is a use case of using ontology in E-government (i.e., data and process governance) to tackle the issues of semantic and organizational interoperability. However, the Palestinian e-Government had Interoperability issues related to heterogeneous Information Systems.

Even this challenge to achieve interoperability in E-Government is also focused by Fernando Ortiz-Rodríguez, Raul Palma, Boris Villazón-Terrazas [17] who have presented EgoIR, an ontology-based information retrieval system intended for E-Government which aims to retrieve the government documentation from government back office to allow final users (citizen and business) to get what they need in any time and any place. In their similar efforts Ortiz-Rodriguez [13] once again used a set of government ontologies to represent Mexican local government processes. They presented the results obtained in an ongoing project commissioned by the Mexican government that seeks strategies for the e-Government to reduce the complications faced when delivering services to citizens.

Another benefit of semantic web and ontologies in E-Government sector is reusability which is achieved by Dr. Mohammed T. Al-Sudairy and T. G. K Vasista from Riyadh, KSA [10] who enable proper integration of knowledge in a way that is reusable by several applications across governance business from discovery to ministry affairs. Also Jean Vincent Fonou Dombeu [14] talked of constructing ontology in support of e-government adoption processes in Sub Saharan Africa by applying a five step framework to methodically gather concepts and activities of the domain and to build the ontology. It is worth noticing in the work of H. A Santoso, Z. T Abdul-Mehdi, S Haw [11] that OntoGov

Project is developed as a semantic platform for composition, configuration and evolution of e-Gov service; SmartGov as a knowledge-based platform for assisting online transaction on public sector employee, in passing [11].
Graciela Brusa, Ma. Laura Caliusco, Omar Chiotti [18] have presented the ontology building in a local public sector: the budgetary and financial system of Santa Fe Province (Argentine) and have depicted the information integration in E-government domain.

The literature discloses that many e-government projects in developed countries use ontology for data type description and web service features, and that the very few current e-government applications that are being run in India lack knowledge base components as ontologies that support their sustainability. We believe that this situation is in part due to so much diversity in Indian subcontinent and the lack of proper protocols of collaboration between industries, governments and current research efforts towards promoting ICT and e-government along with Ontology implementation in the country.

IV. Semantics In E-Government Sector

In the past several decades, we have seen an enormous number of heterogeneous and often autonomous information systems related to corporations, governmental institutes, financial institutions, universities, health institutes or hospitals, etc. There is a growing demand to integrate the information systems to achieve reusability, knowledge sharing and provide more seamless services. Though, the interoperation of heterogeneous and autonomous organizational information systems is a major challenge. It is even more stimulating in case of a government i.e., to construct an e-government because of the complexity, diversity, and multiplicity of public sector institutions.
For instance, often, each sector in the government forms an independent government on its own and interoperation among the departments of one sector is not found. In fact, the organizational complexity of the government and its various sectors results in more heterogeneous 'islands' of information systems each of which is not interoperable[9].
Hence the most important need is to add the isolated applications of different departments and provide a single face to the end user, who should not be concerned with aggravations when performing government activities and to achieve the final outcome as desired.

Hence the problems faced by the both Government and citizens individually are as under:

a) Problem of data exchange between heterogeneous information systems.
b) Lack of any common ground upon which integration of various systems is possible.
c) Absence of interoperability between various individual government departments.
d) Redundancy of each citizen data at every individual government departments.
e) Problem of use of different names, structures or scales for the same kind of information.

Semantic web and Ontologies give a highly dynamic and flexible plan of the information within a domain. The Ontologies generation is an important activity to enable semantic data integration. The construction of ontology helps various applications that delivers the correct information at the correct time to make better informed decisions in applications such as marketing, sales, public policy etc., in passing. Thus a range of semantic technologies based on ontologies enable the proper integration of knowledge that is reusable by variety of applications across governance or business explicitly as in [12].

A. Issues and Challenges in Indian E-Government Perspective

The present implementations of the Indian Government websites are based on Web 1.0 and Web 2.0. There is a lot of data that is distributed and spread across various websites. There is a need to define methods by which we can use technology to bring the metadata and content related to them under one unit while at the same time give the end user the freedom to use data as a de-centralized separate unit too. But when such systems get implemented, the size of the data managed as a single repository for the entire country become very large. The benefits derived from such systems can be increased if data is stored on the websites in such a way that is understood by machines and processing happens without human intervention.

Looking at the current web based implementations, it is clear that while the base of all government websites is still the same, but they show different aspects of the Indian government to the citizens. These multiple websites have their own databases and data processing logics and all these websites exist in singleton and are not connected to each other.
The current implementation of the Indian government websites have issues/ problems defined as under:

1. Data inconsistency across multiple websites

Similar data needs to be searched but there is a confusion which website to consider, state or national level as there is overlapping in nature. This leads citizens' in a confusing state of mind as to which websites need to be accessed for the corresponding information. Also another important factor is, will the data be saved in both the databases individually or mutually by the two different websites? How will they be updated or managed?

Example, If an end user wants to search information related to the financial budget for Hyderabad, the user can be confused on whether to access the Ministry of Finance

website of the Indian Government or check the website of Andhra Pradesh.

2. *Insufficient support for the Information explosion*

The current government websites are informational based websites having abundance of information available to the citizens for viewing. As the information is increasing tremendously, it is evident that a better method for storage and retrieval needs to be incorporated that keeps pace with the latest trends of World Wide Web also.

3. *Available information not in sync with automated machine processing*

The information found on the web is in HTML format which is readable and explicable by humans. But sadly, it does not match with the format that can be understood by machines for any kind of further processing. Hence, though there is plenty of information available on the website, it cannot be used by many applications automatically for any processing. Some sort of manual working is required for extracting data at such places and then using it in other applications.

4. *Data redundancy between various government departments*

This is a common problem found among various government departments like Electricity department, Gas Authority, Banking sector, Ration card, etc. that multiple records corresponding to one and same person is recorded at various departments individually creating duplication of data.

CASE 1: Redundancy of personal details of a single person at every government department corresponding to his account, here every account is a single entity though.

For example, a person named "Mukesh" has account in electricity department for generation of electricity bill. He again has his details recorded in his local Gas agency to avail the facility of LPG cylinders. Next, Mukesh also has a ration card which again stores his personal details like name, address, contact details, family dependents, etc. all stored again and again at multiple locations creating data redundancy.

CASE 2: Redundancy of personal details and duplication of connections at every government department corresponding to multiple accounts for a particular House No.

This refers to a case where a single house depicts multiple connections of electricity bill or gas agency accounts at same residence using separate entities. For example, for a given house address "3/30, Gomtinagar", there are two electricity connections and two different gas connections are found registered under two different names, thereby showing mesh of redundant data and misuse of government resources.

For all the problems mentioned above, Semantic Web Technologies and implementing Ontologies looks like one of the most favorable answers. The use of Semantic Web Ontologies for web implementations in E-Government Sector will benefit both end users and the decision making in the government [15].

There is a lot of scope for work in the government domain. The implementation of Semantic Web ontologies is at a very early stage in the Indian context and there is a massive scope for implementation which would make the data and content related to the government readily accessible to citizens. This would also in the future help in providing better analysis tools to the government for better decision making [15].

B. *Solution for the Issues*

In future we like to project a useful online service for the citizen of India which aims to bridge the gaps between the Indian Government and its citizens by identifying the workability of various departments of Govt. Of India, all of which act as independent units in their own way.

Ontology based information retrieval will play an important role; thereby bridging the gap between Government of India and the fellow citizens and giving them a common platform for automatic information retrieval from multi-faceted environment of Government sector.

We propose building ontology (See Fig 1) in following two phases: Firstly design semantic-enhanced knowledge-base that will help to integrate the different state government agencies in one place. For this, we need to list various conceptual definitions and define links between them; secondly we will construct ontology, i.e., implementation of above knowledge base using a formal tool like Protégé, and then verify the consistency, populate the given ontology, and finally draw inferences on the concepts and instances.

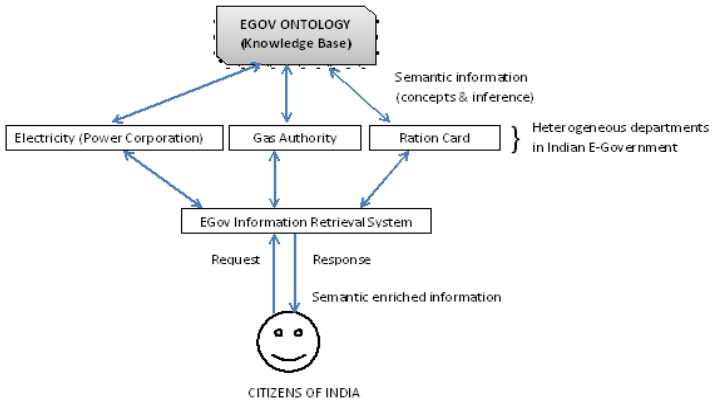

Fig 1: Semantically enriched Information Retrieval System

V. CONCLUSION

In the paper we have discussed various issues and problems we find in manual Indian government and workability of E-Government. We have focused on many situations where efficient information retrieval was not possible. We even saw how Semantic Web ontologies can very easily become

solution to all above problems. An efficient automatic information retrieval system needs to be devised that could cater to all mentioned problems related to lack of interoperability, heterogeneous and island departments and give an opportunity to citizens, businesses and Government to integrate and recover ontologically sound information.

VI. FUTURE SCOPE

We propose to amalgamate heterogeneous knowledge taken from various government sources as one unit and decode unique knowledge and issue various queries that will help to predict the future patterns and necessary information that can be beneficial both at government as well as citizens' end.

REFERENCES

[1] Borst, W.N.: Construction of Engineering Ontologies. University of Tweenty. Enschede, NL-Centre for Telematica and Information Technology. (1997)

[2] T. Berners-Lee, J. Hendler, and O. Lassila, "The Semantic Web", Scientific Am, May 2001, pp. 34– 43.

[3] A. Barbagallo, A. De Nicola, and M. Missikoff, "eGovernment Ontologies: Social Participation in Building and Evolution", in the 43rd Hawaii International Conference on System Science, 2010, Proc. IEEE, pp.1-10.

[4] Gruber, T.R.: Toward Principles for the Design of Ontologies Used for Knowledge Sharing. Presented at the Padua workshop on Formal Ontology, to appear in an edited collection by Nicola Guarino (1993)

[5]] Holsapple, C. W., Joshi, K. D. A collaborative approach to ontology design, In Communications of the ACM, 45(2), 42-47, 2002.

[6] Neches, R., Fikes, R.E., Finin, T., Gruber, T.R., Senator, T., Swartout, W.R.: Enabling Technology for Knowledge Sharing. AI Magazine. 12(3) (1991) 36-56.

[7] D. Fensel, "Ontologies: Silver Bullet for Knowledge Management and Electronic Commerce", Springer, 2001.

[8] Studer, R., Benjamins, V.R., Fensel, D.: Knowledge Engineering: Principles and Methods. Data and Knowledge Engineering. 25: 161-197. (1998)

[9] Mustafa Jarrar, Anton Deik, Bilal Farraj, "Ontology-based Data and Process Governance Framework-The Case of e-Government Interoperability in Palestine".

[10] Dr. Mohammed T. Al-Sudairy and T. G. K Vasista, "Semantic Data Integration Approaches For E-Governance", International Journal of Web & Semantic technology (IJWesT) Vol.2, No.1, January 2011.

[11] Santoso H. A., Abdul-Mehdi Z. T., Haw S. (2009) "Semantic Enhancement Framework for e-Government Using Ontology Versioning Approach", The 6th International Conference on Information Technology and Applications (ICITA 9-12-, Nov. 2009), Hanoi, Vietnam, ISBN:978-981-08-3029-8.

[12] Gardner S. P. (2005) "Ontologies and Semantic Data Integration", Drug Discovery Today, Vol.10, Issue 14, p1001-1007.

[13] F. Ortiz, Rodríguez, Mexican e-government ontologies: an adaptation, Fourth LACCEI International Latin American and Caribbean Conference for Engineering and Technology (LACCET'2006), Mayagüez, Puerto Rico, 21-23 June, 2006.

[14] Jean Vincent Fonou Dombeu, "A Conceptual Ontology for E-government Monitoring of Development Projects in Sub Saharan Africa", IST-Africa 2010 Conference Proceedings, 2010.

[15] Mr.Bonson Sebastian Mampilli, J. Meenakumari, "A Study On Enhancing E-Governance Applications Through Semantic Web Technologies", International Journal of Web Technology, 2012.

[16] D. Tidwell, "Web Services-The Web's Next Revolution", IBM Web Service Tutorial, 29 Nov. 2000, http://www-106.ibm.com/developerworks/edu/ws-dwwsbasics-i.html.

[17] Fernando Ortiz-Rodríguez, Raul Palma, Boris Villazón-Terrazas, "EgoIR: ontology-based information retrieval intended for eGovernment".

[18] Graciela Brusa1, Ma. Laura Caliusco2, Omar Chiotti3, "Building Ontology in Public Administration: A Case Study".

[19] M. Paolucci, T. Kawamura, T.R. Payne, and K.P. Sycara, "Semantic Matching of Web Services Capabilities", In Proceedings of International Semantic Web Conference, 2002, pp.333-347.