

Towards Metadata and Data Standards for Semantic E-Governance Services in India

Devika P. Madalli, Indian Statistical Institute, 8th Mile, Mysore Road, R.V. College Post, Bangalore 560059, Karnataka, India; E-mail: devika@drtc.isibang.ac.in

ABSTRACT

E-Governance facilitates outreach of government services to the citizens with help of information and communication technology. Metadata has been considered as an essential element in any digital environment. The paper discusses the Indian e-Governance efforts and enlists its different initiatives. The role of MetaData and Data Standards (MDDS) in e-governance is highlighted. Need for Interoperability in e-Governance is also discussed. A conceptual model towards achieving MDDS is presented.

INTRODUCTION

Internet has a marked impact on the methods of communication, business processes, commerce, research and academics as also in governance. Though it was primarily the businesses and also academics and research communities that initially adapted Internet into their activities, governments have also realized the importance of online reach to citizen and end users [Butt and Persaud, 2005]. One of the chief challenges for government departments is to outreach their programs to the wide spread communities considering physical aspects, such as distance, weather, terrains etc. E-mode of operation presents a viable and efficient option to governments to deliver services to the different sectors they deal with, overcoming the physical constraints. Also a larger mass of the end-users can be reached through e-governance systems. The use of Internet technologies has improved the effectiveness and efficiency of governance services [Agarkhani, 2003].

E-governance has different facets to it, in the sense that the models and the systems vary according to department and its services offered. Many e-governance systems - some domain specific and others task specific - have been discussed [Rotchanakitumnuai and Speece, 2003], (Sathye, 1999), (Doherty; Ellis-Chadwick and Hart, 2003), (Levy and Powell, 2003)]. The e-governance models depend on constitutional, cultural, economic fabrics of the government offering the services. This is evident in the differences that emerge between models adapted by different countries for their e-governance initiatives. [UKGovTALKa, AGLS].

In India, the National eGovernance initiative was started by the Government of India (GOI), with a vision to offer faster, improved and more efficient services, shared resources and services, increased productivity and to establish standardized systems and processes [NICa] to manage government services and serve patrons of different sectors.

1. E-GOVERNANCE IN INDIA

With a wide vision mentioned above and objective of outreaching government services to public, Indian government has started the e-Governance initiative. The main problems with individual systems for e-Governance developed in India so far is that there is too much data and not enough information. In the sense that there are no cohesive and meaningful services though a lot of data is available. Also the systems were developed independently as stand-alone with no common data architecture and are too expensive to bridge [NICa]. GOI aims to address these issues through the National E-governance Action Plan (2003-07). A few initiatives taken in this regard are as follows:

- Initiative on Digital Signature
- Initiative on EG Standard
- Initiative on India Portal

- Initiative on India Development Gateway (InDG)
- Initiative on National e-Governance Service Delivery Gateway (NSDG)
- Initiative on Replication
- Initiative on Technology Development for Indian Languages (TDIL)
- Initiative on EG Capacity Building
- I T Act & Rules
- Assessment Framework
- State Data and Service Centres

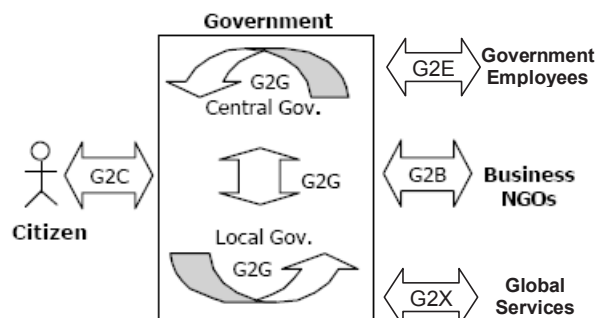
The work is taken up by National Informatics Centre (NIC), Department of Information Technology under the Ministry of Communication & Information Technology.

2. LEVELS OF INTERACTION IN E-GOVERNANCE

The level of e-Government services, NIC has defined, includes five main key players -Government, Citizen and Business, Employees, and Global Services. The most likely levels of interaction in this context can be: G2C, G2B, G2G, G2E and G2X.

- **Government-to-Citizen (G2C):** This is the interaction between Government and citizen. In this interaction, the citizens may utilize the information provided by the government at the service portal or gateway; sends email to concern government department, etc. For e.g. the farmer can check out the latest market price of wheat at the information portal containing latest crop prices.
- **Government-to-Business (G2B):** At this level of interaction, the business personnel can fill tenders, pay sales tax, enquire about the latest company laws, etc online to the concerned government department.
- **Government-to-Government (G2G):** This is the mode of interaction between two or more government departments for the smooth accomplishment of government functioning.
- **Government-to-Employee (G2E):** This is the mode of transaction between government to employee (Govt.) for accessing their pay slip, and other related

Figure 1. Interactions in e-governance [National e-Governance Plan, 2005]



transactions with the employer (i.e. respective Govt. Department). “Daily” is an intranet site of the Department of Rural Development, Govt. of India, promoting G2E governance within the ministry. It provides an easy, efficient and one-stop access to all information and services required by the employees to carry out their daily work [NIC, 2005].

- **Government-to-Global Services (G2X):** This is the mode of transaction between Government to various global services implemented and executed within the country or abroad.

In all the ways of communication at the levels indicated above, it is important that different stakeholders must have the ability to exchange and mutually use information. This is only possible when information is represented in meaningful manner with indicative tags. This is the function of metadata in e-governance.

3. METADATA

The most common definition of the term ‘metadata’ is structured data about data, information that describes other information. The World Wide Web Consortium has defined it as: “*machine understandable information for the Web*”. For example, if a Web page has an author, a title, a date of creation and a unique Internet address, these elements constitute metadata about the page. Metadata is an Internet- age term for information that librarians traditionally have put into catalogs and it commonly refers to descriptive information about Web resources [Ercegovac, 1999].

A metadata record is a label, like the label we might find on an everyday product in a supermarket. The label describes the product, and quite often contains information that is useful both to consumers and to the electronic systems that control the store. The label contains data a person can understand, such as a list of ingredients, and data a machine can read, like a bar code [UKGovTalk, 2006].

Metadata labels are attached to documents and other information assets with similar aims in mind. By storing information such as author’s name, version number, subject, we can help people to find information more easily, and allow computers to process it more effectively.

4. ROLE OF METADATA IN E-GOVERNANCE [EGMS, 2004]

Metadata standard can be used to classify and categorize Government information and services, facilitating identification of services and information intelligently. Inter-departmental information exchange becomes easier. Hence, metadata increases the visibility and accessibility of Government services over the Internet [Weibel, 2005]. Modernizing Government calls for better use of official information, joined-up systems and policies, and services designed around the needs of citizens. Metadata makes it easier to manage or find information [EURIM, 2002], be it in the form of web pages, electronic documents, paper files, databases, anything. For metadata to be effective it needs to be structured and consistent across organizations.

However while metadata aims to represent what data it carries, it is also important that the data itself should be represented following international standards in order to achieve interoperability between systems.

5. DATA STANDARDS

It can be said as, agreed upon terms for defining and sharing data. According to ISO, a standard is “documented agreements containing technical specifications or other precise criteria to be used consistently as rules, guidelines, or definitions of characteristics, to ensure that materials, products, processes and services are fit for their purpose.” There are several types of data standards including terminology standards and message format standards [PHDSC, 2006].

Some important data standards which are commonly used in e-Governance services are:

- Author name format
- Personal information
- Address information
- Date/Time format
- Organizational information
- Financial Information; among others

6. INTEROPERABILITY ISSUES

Interoperability means the ability of information and communication technology (ICT) systems, as well as, of the business processes they support in order to exchange data and enable the sharing of information and knowledge [Edutools Wiki].

6.1 Why Interoperability

Effective e-Government results in reducing the amount of paperwork, streamlining processes and considerably enhancing citizen services and government communication. There are indeed many reasons that motivate governments in e-Government implementation, e.g.:

1. Reduction of expenditure: reducing time spent on administration, thus resulting in cost savings;
2. Improvement of professionalism in agencies management and modernization of the internal organization: e-Government will make internal government processes and external interaction more efficient;
3. Improvement of the quality of service and customer satisfaction.

It is ideal that each government department’s information is encoded in the same format and also the services are planned in similar layers of data representation, service model and interface layers. But often this is not possible in practical governance. Each department has its particular needs in terms of data sets and also the expected forms of interaction and ultimate delivery of services. Hence, the e-governance method varies from department to department. Also, another reason for disparity arises because different department have moved to e-governance at different times and follow the standard and technology available at that time. Added to this is the disparity of cultural influences in approach to information and hence its representation. Another level applicable to diverse nation such as India is the multi-lingual communities that warrants cross lingual retrieval. This particular issue however, is outside the scope of this paper.

All of the above reasons warrant that measures of interoperability be implemented in e-governance system.

6.2 Crosswalk

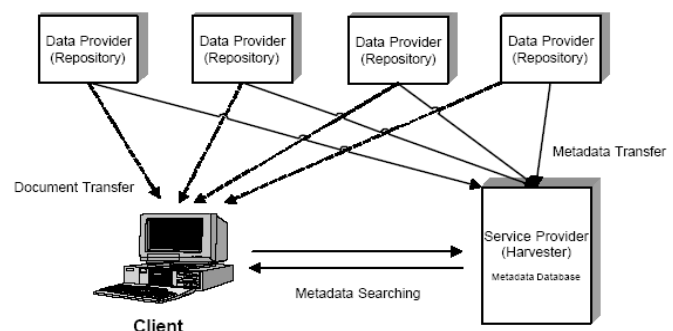
Crosswalks or metadata mapping support the ability of search engines to search effectively across heterogeneous databases, i.e. crosswalk helps to promote interoperability [LAOAP]. Crosswalk may be achieved through a table that maps the relationships and equivalencies between two or more metadata formats. This is rather simply said than possible, because of the wide disparities in metadata sets and data representations in practice.

6.3 Harvesting Technology

Harvesting refers to the gathering together of metadata from a number of distributed repositories into a combined data store [OAI, 2002].

One of the most interesting recent developments in the digital library arena involves the Open Archives Initiative (OAI). The Open Archives Initiative emerged out of the scholarly communications arena as a means to provide interoperability among multiple information sources. The communications model of OAI relies

Figure 2. OAI-PMH architecture



on the bulk transfer, or harvesting, of metadata between a service and all its data providers, based on a set of very simple protocols.

The OAI universe is based on information repositories, or “data providers,” that make their metadata available, using a prescribed set of protocols, to “service providers” that build new information resources. End-users gain the benefit of OAI-based services that aggregate the metadata of multiple OAI repositories. It is to be noted that OAI operates with metadata, not complete works of digital content. In most cases the metadata include links back to the original information repositories for access to the documents or other digital objects [Mittal, Kumar, Mohania, Nair, Batra, Roy, Saronwala, and Yagnik, 2004].

7. ROADMAP TO INDIAN E-GOVERNANCE METADATA AND DATA STANDARD

Steps for developing metadata and data standards for e-governance are as follows:

1. **Evaluation of international metadata standards:** Some major e-governance standards has to be studied and evaluated to get the idea about the basic elements, metadata descriptors, vocabulary control and metadata management mechanism for development of metadata standard
 - a. e-GMS
 - b. GILS
 - c. AGLS
 - d. NZGLS

2. **Identification of metadata elements for Indian context:** Selection of basic metadata element sets and in conformity with the world standard for e-governance. As well as various application profiles suitable for the needs of various domains like agriculture, finance and economy, etc has to be developed.
3. **Controlled vocabulary:** Each government sector has its own vocabulary regarding its function and subject area. Therefore, two controlled vocabulary lists have to be prepared: one dealing with the services and another with the subject terminology.
4. **XML schema and XSLT for domain specific services:** To start web-based services, the domain specific service description has to be represented in the form of XML/RDF/OWL (Ontology). Its presentation on the web browser has to be taken care by their XSLT.
5. **Data standards:** The data which are to be input in the respective metadata fields also require standardization in terms of author name format, date format, organizational information, financial information etc.
6. **Metadata interoperability framework:** This framework is to ensure the interoperability between various services from diverse domains.

7.1 Use Case based on Conceptual Model for Metadata and Data Standard

The conceptual model has been framed to explain the role of metadata description in e-governance services. This also shows the usability of data standards in the given context. For example, a citizen looking for the information on agriculture can interact with the web portal providing the web services. The web service providing the required information is based on the interoperability framework. This interoperability framework works to translate the content of two different databases using different content representation format into the desired format. In databases, content is described using the e-government metadata elements in form of ontologies. The values of each of these elements can be standardized with pre-structured data standards. One of the instances is a farmer's enquiry for the best prices for his commodity. Ideally the query should return the best priced market with nearest reach. It combines departmental information about agricultural commodity, related markets, and other information like weather, transport etc. The interoperability is achieved through representation or description layers that support the discovery layer as shown in the conceptual model below:

8. CONCLUSION

e-Governance is quite complex to achieve with variance in the communities, services and formats among several other factors. Metadata plays an important role in providing a proper and standard representation of the e-governance metadata. But in many e-governance services and departments where data is different, there would be need for different sets of metadata elements for comprehensive representation. Yet other scenario is that different departments started e-governance with different data sets, standards and technological tools. This brings to for the problem of variation in standards and the only way to bridge this kind of

Figure 3. Roadmap to metadata standard for e-governance in India

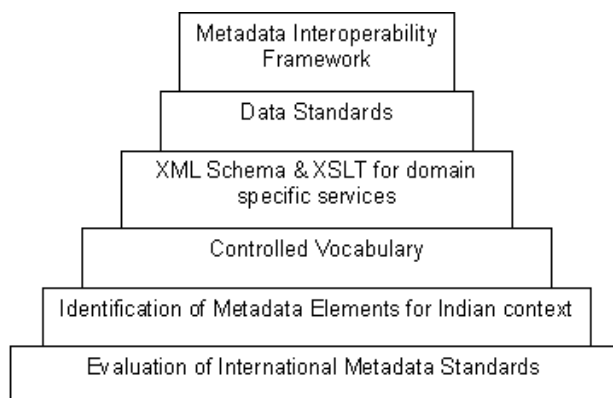
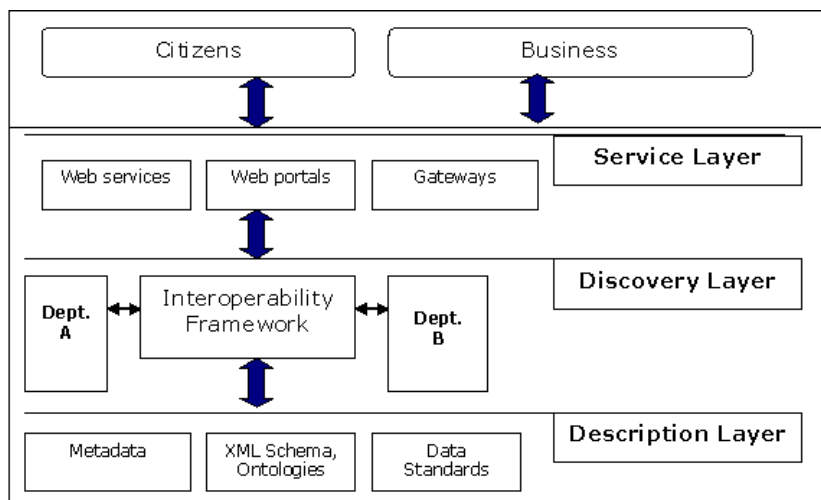


Figure 4. Conceptual model of interoperability framework for e-governance in India



disparity is to build crosswalks so that even when the end user approaches are different for the same data, it is still retrieved. Further, true interoperability can only be achieved when semantic interoperability can be achieved. A model, as illustrated, that is based on ontologies along with the metadata and standard data representation, in the description layer, would facilitate the discovery layer in vending services to patrons in varied scenarios.

9. REFERENCES

1. AGLS metadata element set. http://www.naa.gov.au/recordkeeping/gov_online/agls/metadata_element_set.html
2. Asgarkhani, M. (2003). A strategic framework for electronic government. Proceedings of the 22nd National IT Conference, Sri Lanka, pp57-65.
3. Butt, Irfan and Persuad, Ajax (2005). Towards a citizen centric model of e-Government adoption. Proceedings of the 3rd International Conference on e-Governance, Lahore, Pakistan, pp6-15.
4. Doherty, Neil; Ellis-Chadwick, Fiona and Hart, Cathy (2003). An analysis of the factors affecting the adoption of the Internet in the UK retail sector," Journal of Business Research, vol. 56, no. 11, pp. 887-897.
5. eGMS (e-Government Metadata Standard), Version 3.0, (2004). <http://www.govtalk.gov.uk/documents/eGovMetadataStandard%2020040429.pdf>
6. Edutools Wiki. Glossary Analysis. <http://www.edtechpost.ca/pmwiki/pmwiki.php/Main/GlossaryAnalysis>
7. Ercegovac, Z. (1999). Introduction. Journal of the American Society for Information Science, Vol. 50, no. 13, 1165- 1168.
8. LAOAP. Latin America Open Archives Portal- glossary of terms. <http://lanic.utexas.edu/project/laoap/glossary.html>
9. Levy, Margi and Powell, Philip (2003). Exploring SME Internet Adoption: Towards a Contingent Model," Electronic Markets, vol. 13, no. 2, pp173-181.
10. Mittal, P. A.; Kumar, M.; Mohania, M. K.; Nair, M.; Batra, N.; Roy, P.; Saronwala, A.; and Yagnik, L. (2004). A framework for eGovernance solutions. IBM Journal of research and development. Vol. 48, no. 5/6. <http://www.research.ibm.com/journal/rd/485/mittal.html>
11. National e-Governance Plan (2005). Approach and key components. <http://www.mit.gov.in/plan/NGEP/NEGP11-03-05/NEGP%20WS%20Pres%20-%2011%20Mar%202005.ppt>
12. NIC (National Informatics Centre), (2005). Good governance through ICT. NIC: New Delhi. 2005.
13. NICa. Metadata and data standards for e--Governance Applications. http://egovstandards.gov.in/standards_metadata_app
14. OAI (Open Archives Initiatives) (2002). Information Today. <http://www.librarytechnology.org/litg-displaytext.pl?RC=9627>
15. PHDSC (Public Health Data Standards Consortium) (2006). Public Health Data Standards Tutorial - Glossary of Terms. www.phdatastandards.info/knownresources/tutorials/glossary.htm
16. UKGovTalk (2006). A layman's introduction to metadata www.govtalk.gov.uk/documents/Laymans_guide_to_metadata%20v1.1.doc
17. EURIM (2002). Interoperability- Joined Up Government Needs Joined Up Systems. The European Information Society Group, Briefing No 36. www.eurim.org.uk/resources/briefings/br36.pdf
18. UKGovTALKa. GovTalk- Information on policies and standards for e-government. <http://www.govtalk.gov.uk/>
19. Weibel, Stuart L. (2005). Reflections on a decade of metadata consensus building. D-Lib Magazine. Vol. 11, no. 7/8.

0 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/proceeding-paper/towards-metadata-data-standards-semantic/33278

Related Content

Trend-Aware Data Imputation Based on Generative Adversarial Network for Time Series

Han Li, Zhenxiong Liu, Jixiang Niu, Zhongguo Yang and Sikandar Ali (2023). *International Journal of Information Technologies and Systems Approach* (pp. 1-17).

www.irma-international.org/article/trend-aware-data-imputation-based-on-generative-adversarial-network-for-time-series/325212

The Internet of Things

Nigel McKelvey, Kevin Curran and Nadarajah Subaginy (2015). *Encyclopedia of Information Science and Technology, Third Edition* (pp. 5777-5783).

www.irma-international.org/chapter/internet-things/113032

Lean Logistics of the Transportation of Fresh Fruit Bunches (FFB) in the Palm Oil Industry

Cheah Cheng Teik and Yudi Fernando (2018). *Encyclopedia of Information Science and Technology, Fourth Edition* (pp. 5422-5432).

www.irma-international.org/chapter/lean-logistics-of-the-transportation-of-fresh-fruit-bunches-ffb-in-the-palm-oil-industry/184245

Classification of Sentiment of Reviews using Supervised Machine Learning Techniques

Abinash Tripathy and Santanu Kumar Rath (2017). *International Journal of Rough Sets and Data Analysis* (pp. 56-74).

www.irma-international.org/article/classification-of-sentiment-of-reviews-using-supervised-machine-learning-techniques/169174

Modeling Knowledge Society

Lech W. Zacher (2015). *Encyclopedia of Information Science and Technology, Third Edition* (pp. 7192-7201).

www.irma-international.org/chapter/modeling-knowledge-society/112417