The Governmental Markup Language (GovML)

The Governmental Markup Language (GovML)

Efthimios Tambouris[(1)], Gregory Kavadias[(2)] and Elias Spanos[(2)]

[(1)]Independent Consultant, Agion Anargyron 28, 175 64, Athens, Greece. tambouris@iname.com

[(2)]Archetypon S.A., Sygrou Ave. 236, Kallithea, 176 72 Athens, Greece. {gkavadias, espanos}@archetypon.gr

Author Bio: Dr. Efthimios Tambouris is an independent consultant. Before that, he was founder and manager of the eGovernment unit at Archetypon S.A. and researcher at the Hellenic National Center for Scientific Research "Demokritos". He holds a Dipl-Eng from National Technical University of Athens, Greece and an M.Sc. and Ph.D. from Brunel University, UK. He initiated and/or managed several e-government projects (IST EURO-CITI, eGOV, SmartGov, OntoGov) and serves as expert at the European Committee for Standardization (CEN). His research interests are in the areas of services to citizens (mainly e-government and e-democracy) and high performance computing.
Author Bio: Mr. Grigoris Kavadias received his B.Sc. and M.Sc. in Computer Science from the University of Athens, Greece. Since 2000 he works with the Greek Company Archetypon SA for several research and commercial projects related to the area of e-government e.g. IST eGOV (www.egov-project.org), SmartGov (www.smartgov-project.org) and OntoGov (www.ontogov.com). His research interests are in the areas of mobile/pervasive computing, semantic web (e.g. RDF, ontology languages) web services and relational databases. He is an IBM certified specialist: IBM DB2: "Family fundamentals". He is the author of several publications in the area of e-government.
Author Bio: Mr. Elias Spanos joined Archetypon S.A. in mid 2000 as a Project Manager also involved in the analysis and deployment of enterprise and governmental IT systems. Recent activities include the technical coordination and management of projects especially in the e-government and e-democracy areas. He has been the scientific coordinator in the EC co-funded research projects IST EURO-CITI and eGOV. His scientific areas of interest include portals, IT middleware platforms and metadata. Research on these areas is reported through several publications in conferences and journals. Mr Spanos is a member of the Technical Chamber of Greece.

The Governmental Markup Language (GovML)

Efthimios Tambouris, Gregory Kavadias and Elias Spanos

The value of public sector information can be significantly increased by employing standards. For example, the use of a metadata standard for describing governmental resources allows citizens and businesses to locate and retrieve these resources in a more efficient and effective way. Furthermore, the use of a common document template for describing elementary and integrated public services (such as life events) can also substantially improve management and reuse of information and increase citizens' satisfaction. This paper proposes a metadata element set for governmental resources and a template for describing life events and pubic services. The proposed metadata element set can be used to facilitate the search and retrieval of governmental resources, such as electronic documents. The proposed document template, termed the Governmental Markup Language (GovML), is a recommendation for describing public services and life events. It is anticipated that both public organizations and consumers of public services (citizens, businesses and other public organizations) would benefit from such a common information structure. Public authorities will be able to better manage (create, modify, store etc), exchange and reuse their information. Citizens and businesses will enjoy better services when searching for information regarding public services and life events. From a more technical perspective, GovML data vocabularies are proposed XML patterns, validated against an appropriate XML Schema. Interoperability among public authorities is ensured as the rules that dominate the document syntax are common. Moreover, GovML-compliant information exploits XML potential thus can be displayed in multiple formats and devices, such as web browsers, PDAs and mobile phones.

e-government, XML, GovML, generic description, specific description, public services

INTRODUCTION

Governments all over the world start realizing that within the exercise of its public tasks, the public sector collects, processes and disseminates huge quantities of information (European Commission, 2002). In Europe, a report by eContent (2000) revealed that "public bodies are by far the largest producers of information". The European Commission estimates the value of public sector information at 68 billion Euro per year and recognizes its potential to boost economic activity and job creation (eContent, 2000; European Commission, 2001). At the political level, a Directive of the European Parliament and the Council on the re-use and commercial exploitation of public documents has been recently published (European Commission, 2002).

It is now recognized that employing a metadata standard is essential for the discovery and retrieval of governmental resources. Metadata is commonly defined as "data about data" or "information about information". The use of a metadata standard allows describing resources, thus enables their discovery and retrieval by public authorities, citizens and businesses.

An important part of public sector information is descriptions of public services. Indeed, providing proper descriptions of public services allows citizens and businesses to be better informed, thus reduces frustration when interacting with the public sector. It should be noted that descriptions of public services are often available in the Web sites of public authorities, even in the case where these services are not provided electronically. Examples include national portals, such as www.firstgov.gov (USA), www.direct.gov.uk (UK), www.fonction-publique.gouv.fr (France), www.help.gv.at (Austria), and www.ecitizen.gov.sg (Singapure). All these portals provide descriptions about public services even when the services themselves are not provided online. Furthermore, some of these portals provided descriptions about integrated public services. It is common for example that public services are integrated around life events for citizens and around business episodes for businesses. In all cases, quality of relevant information would be greatly enhanced by using common templates for describing life events and public services.

The aim of this paper is two-fold. Firstly, to propose a metadata element set for governmental resources. Secondly, to propose the Governmental Markup Language (GovML), as a recommended set of common templates for describing life events and public services.

Both the metadata element set and GovML are results of research that was conducted within the IST eGOV project (www.egov-project.org) aiming to specify, develop and evaluate an integrated platform for integrated service delivery based on life events (Tambouris E. and Wimmer M., 2004; Tambouris E. and Spanos E., 2002). The introduction of GovML has been recently recognized by the European Commission as a significant contribution to interoperability (European Commission, 2003).

TOWARDS IDENTIFYING GovML ATTRIBUTES

Introduction

Public authorities worldwide are increasingly using the Web in order to inform citizens and businesses on their services. This online information however is rarely compliant to a standard structure. On the contrary, it is common that each public authority applies its own format to the presentation of information, hence preventing users from a unique experience whenever interacting with the public sector. This is usually the case even providing information on the same service by

different authorities.
Lately, public authorities are also providing online information on integrated public services to better cover the needs of their citizens and businesses. A useful metaphor for integrating public services for citizens is that of "life events", such as "getting married". Here, public authorities structure their Web content around life events thus creating a more user-friendly Web environment for citizens. Even in this case however, the information on life events does not usually follow a common structure.
In this section, a set of common templates for describing life events and public services, termed Governmental Markup Language (GovML), is proposed.
The use of a common set of templates and storing formats by public authorities has a number of advantages for both public organizations and consumers of public services (citizens, businesses and other public organizations). For example, public authorities will be able to better manage (create, modify, store etc), exchange and reuse their information. Furthermore, citizens and businesses will enjoy better services when searching for information regarding public services and life events.

Methodology

Te proposed Governmental Markup Language (GovML) is a recommended set of common templates for describing life events and public services (Kavadias G. and Tambouris E., 2003). GovML was specified within the IST eGOV project by the consortium partners. More specifically, the eGOV consortium consisted of 10 partners coming from Austria, Finland, Germany, Greece and Switzerland. Partners represented a mixture of private IT companies, academic and research institutes, and public administrations. The steps that were followed towards the specification of GovML data elements are (Figure 1):

1 All partners performed an analysis of governmental portals and Web sites. The analysis focused on the structure and presentation of the public sector content from the users' point of view.
2 A number of public services and life events were investigated in an attempt to identify the necessary elements needed to describe them.
3 The contributions of all partners were assembled and the first version of the proposed GovML data elements was circulated among the consortium.
4 All partners commented on the first version of GovML data elements and proposed refinements. All comments were gathered and evaluated. Thereafter, a new version of the GovML data elements was produced and circulated to the consortium for consultation.
5 Iteration among steps 3 to 5 was realized until consensus was reached between all partners of the eGOV consortium.
6 The final GovML data elements were reported according to ISO/IEC 11179-3 standard (see Kavadias G., Spanos E. and Tambouris E., 2002).

"Insert Figure 1 here"

The final GovML data structure consists of three vocabularies, two for describing public services and one for life events. The three GovML vocabularies are listed in Table 1 and are presented in more details in the next paragraphs.

"Insert Table 1 here"

Generic description data vocabulary for public services

This vocabulary defines a common standard for describing a public service. This information is normally based on a law and can only change when the relevant law changes. This vocabulary is particularly useful for portals that provide general information on public services but they are not themselves service providers. Examples of data elements are: title, procedure, required documents, etc.

Specific description data vocabulary for public services

This vocabulary enables pubic authorities to describe the services that they provide. Here, the values of some elements depend on the public authority that provides the public service. Some of its elements are: name and address of the public authority, public servant contact details, delivery channel of the service, etc.

Data vocabulary for life events and business situations

This vocabulary defines a set of elements necessary to describe life events and business situations. Examples of data elements are: title, description, faq-list, related services, etc.

AN EXAMPLE ON THE USE OF GovML

As an example on the use of GovML data vocabularies the following scenario is assumed. There is a portal that contains information on services provided by local authorities. This portal presents information on elementary services but also presents information on integrated services i.e. life events. It is also assumed that the "getting married" life event has to be described, which consist of a number of services including the service "obtaining a birth certificate".

The portal's content provider may use the "Data vocabulary for life events" in order to present the information on the life event "getting married". More specifically, the content provider will have to include an identifier for this life event, the language in which information is available, the title, a description, those in attention, a faq-list, the related services, and the available delivery channels. The information on "related services" should contain a link to service "obtaining a birth certificate", as this is part of the life event.

The portal's content provider may use the "Generic description data vocabulary for public services" in order to present the information on service "obtaining a birth certificate". Here, all relevant attributes of Table 1 should be filled in with appropriate information. The information that is included here contains everything the citizens need to know about the service. However, as the portal is not itself a service provider (i.e. it does not issue birth certificates) this description does not include details on obtaining the certificate.

On the other hand, local authorities that want to present information on service "obtaining a birth certificate" should use the "Specific description data vocabulary for public services". Indeed, this vocabulary allows providing useful information for consuming the actual service such as contact person details, public authority address etc.

IMPLEMENTATION OF GovML

From a technical point of view, the Governmental Markup Language (GovML) vocabularies are implemented using XML technologies. More specifically, in order to serialize GovML documents in XML format an XML schema was implemented for the validation of their structure. The XML schema validation mechanism (www.w3.org/XML/Schema) was preferred from Document Type Definition (DTD) because it provides a richer set of data-types (including byte, date, integer etc) and allows users to derive their own data types and take advantage of inheritance of elements, attributes and definitions of data-types (Birbeck M. et al. 2001).
New XML documents describing public services or life events can emerge from the XML schema of GovML. Consequently, the appropriate XSL Transformations (www.w3.org/TR/xslt) should be applied for transforming GovML documents to the required format (HTML, WML etc).
XML schema can be easily extended, modified and maintained in the future according to consumer's needs. The full XML schema of GovML can be found in Kavadias G., Spanos E. and Tambouris E. (2002).

EGOV METADATA ELEMENT SET

Introduction

The use of metadata in e-government could enhance citizens' experience of interacting with the public sector using ICT, improve the potential of government to business (G2B) and government to government (G2G) activities and boost economic activity and job creation. As a result, an increasing number of governments are currently launching or maintaining metadata initiatives, which aim amongst others to define appropriate metadata standards for describing governmental resources. At the minimum, the objective of employing metadata is to facilitate the discovery of governmental resources by interesting parties (commonly citizens, businesses and public authorities). However, metadata may be also used to describe resources from various perspectives e.g. to support archiving, to protect intellectual property rights, to manage change of resources etc. It should be noted that governmental resources might be documents, public services, web pages, public authorities or agencies, or anything else e.g. a monument such as Acropolis. In principle, anything can be a resource and it is up to public authorities to select the potential resources for metadata description.

Related Work

The proposed metadata element set is based on Dublin Core Element Set (DC) (www.dublincore.org). DC is a simple metadata element set aiming at facilitating discovery and retrieval of resources. DC is maintained by the DC Metadata Initiative (DCMI), an organization dedicated to promoting the widespread adoption of interoperable metadata standards and developing specialized metadata vocabularies for describing resources that enable more intelligent information discovery systems. Ongoing efforts of DCMI participants include the collaborative development and continual refinement of metadata conventions based on research and feedback between DCMI Working Groups. One of them is the DCMI Government Working Group working aiming to propose additional metadata terms suitable for the e-government domain i.e. for describing governmental information resources.
Today, the DC element set has been used as a basis by many countries when defining their national

metadata standards for governmental resources. The countries that have adopted DC include UK, Ireland, New Zealand, Australia, Canada, Finland etc.
The DC has been also used as a basis from the European Commission's "Interchange of Data between Administrations" (IDA) initiative (europa.eu.int/ISPO/ida/) in an attempt to define a European metadata standard. More specifically, the IDA's project "MIReG: Management Information Resources for eGovernment" (http://europa.eu.int/ISPO/ida/) has specified a metadata standard and implemented a software prototype to be used by Members States for experimentation.
The DC has been also used as a basis from a project by the MMI-DC Workshop (www.cenorm.be/sh/mmi-dc/) of the European Committee for Standardization (CEN) (www.cenorm.be). More specifically, the project "A Dublin Core eGovernment profile and model" has produced the CEN Workshop Agreement (CWA) "Dublin Core eGovernment Application Profiles" that includes a proposed metadata standard based on DC.

Element Set

The aim of the proposed metadata element set is to facilitate discovery and retrieval of Governmental resources across the Internet. The core of the prposed metadata element set is DC Element set. Furthermore, elements from the UK e-GMS framework (www.govtalk.gov.uk/interoperability/metadata.asp?order=title) and Prism (www.prismstandard.org) have been included. Technically, mixing and matching of elements from different sets (such as DC, e-GMS and Prism) is achieved, using mechanisms available in RDF and XML.
The proposed element set was specified and implemented within the IST eGOV project. The methodology used to determine the proposed element set is similar to the one described above for determining the GovML vocabularies. The main objective of the work was to keep the resulting element set as small as possible.
Table 2 presents the eGOV metadata element set.

"Insert Table 2 here"

SUMMARY

In this paper, the Governmental Markup Language is presented as a recommended standard for public authorities to manage descriptions on life events and public services. The use of a comprehensive, standard set of attributes for describing public services and life events has some considerable advantages. It allows public authorities to better manage their content and to provide a better experience to their customers (citizens, businesses, employees or other public authorities). The use of GovML has been recognized by the European Commission as a significant contribution towards achievement interoperability for e-government services.
Furthermore, in this paper a metadata element set for governmental resources is presented. The proposed element set is a recommendation based on Dublin Core, the UK e-GMS and PRISM. The use of metadata in the public sector facilitates discovery and retrieval of Governmental resources across the Internet.

REFERENCES

Birbeck M. et al. 2001. *Professional XML*, USA: Wrox Press Ltd..

eContent, 2000. Commercial exploitation of Europe's public sector information. Retrieved June 1, 2004, from europa.eu.int/information_society/topics/multi/psi/ docs/pdfs/commercial_exploitation/2000_1558_en.pdf

European Commission. 2002. Proposal for a Directive of the European Parliament and of the Council on the re-use and commercial exploitation of public sector documents, COM(2002) 207 final.

European Commission. 2001. eEurope2002. Creating a EU Framework for the Exploitation of Public Sector Information, COM(2001) 607 final.

European Commission. 2003. Linking up Europe: the importance of interoperability for e-government services, Commission Staff Working Document, 16.

Kavadias G., Spanos E., and Tambouris E. 2002. D231. GML syntax and filters implementation. Retrieved June 1, 2004 from http://www.egov-project.org/deliverables.htm.

Kavadias G. and Tambouris E. 2003. GovML: A Markup Language for Describing Public Services and Life Events, *Lecture Notes in Artificial Intelligence, vol. 2645*, 106-115.

Tambouris E. and Spanos E. (2002). An Architecture for Integrated Public Service Delivery based on Life-events, *Electronic Markets, Vol. 12 (4)*, 281-288.

Tambouris E. and Wimmer M. (2004). Online One-Stop Government: A Single Point of Access to Public Services. In Huang W., Siau K. and Wei K.K. (Eds.), Electronic Government Strategies and Implementation (115-145), USA: Idea Group Inc.