

---

# How the US Federal Government is Using XML

An Overview of Selected US Federal Agency Efforts

Kenneth Sall <KSall@SiloSmashers.com>

## Abstract

XML is integral to the US Federal Enterprise Architecture. Its use is actively encouraged by the Office of Management and Budget (OMB) and the E-Government Act of 2002. This paper provides a high-level view of a number of mature XML efforts in the US federal government, including efforts of the CIO Council, E-Government Initiatives, Dept. of Navy (DON), Environmental Protection Agency (EPA), General Services Administration (GSA), Dept. of Justice (DOJ), Small Business Administration (SBA) and others.

## Table of Contents

1. Introduction .....	2
2. Legislative Motivators for XML .....	3
2.1. Clinger-Cohen Act .....	3
2.2. Government Paperwork Elimination Act .....	3
2.3. Section 508 of the Rehabilitation Act .....	3
2.4. H.R. 2458, The E-Government Act of 2002 .....	3
3. Federal Enterprise Architecture and the OMB 300 .....	4
3.1. FEA Reference Models .....	5
3.2. OMB 300 as XML Schema .....	7
3.3. InfoPath Implementation of an OMB 300 E-Form .....	9
4. Working Groups and Federal XML Guidance .....	11
4.1. XML Working Group .....	12
4.2. Web Services Working Group .....	13
4.3. E-Forms for E-Gov Pilot .....	13
4.4. Department of Navy XML Work Group .....	14
4.5. Environmental Protection Agency XML Design Rules .....	14
5. E-Government Initiatives .....	15
5.1. Integrated Acquisitions Environment: GSA .....	16
5.2. E-Travel: GSA .....	17
5.3. Business Gateway: SBA .....	18
5.4. E-Records Management: NARA .....	18
5.5. Recreation One-Stop: DOI .....	18
6. Registry, Repositories and Web Services .....	19
6.1. GSA-NIST Proof-of-Concept XML Registry/Repository .....	19
6.2. GSA Consolidated Component Repository .....	20
6.3. EPA's Environmental Data Registry .....	20
6.4. DoD XML Registry .....	21
6.5. DISA Registry Initiative .....	23
6.6. Web Services and Registries Pilot .....	24
7. Selected Specialized Applications .....	24
7.1. IRS Publications, Instructions, and Guidance in XML .....	24
7.2. House of Representatives XML and Legislative Documents .....	24
7.3. Justice XML Data Dictionary (JXDD) and Justice XML Registry/Repository (JXRR) .....	25
7.4. Generalized Instrument Design System (GIDS): Census Bureau .....	26

7.5. EPA Emergency Response VoiceXML Application for DHS .....	26
7.6. Semantic Web .....	27
8. Summary .....	28
9. Appendix: Early History of Markup Technologies in the US Federal Government .....	28
Bibliography .....	29
Glossary .....	32

# 1. Introduction

An August 2002 General Accounting Office ([GAO](#)) report entitled: *National Preparedness: Technology and Information Sharing Challenges* states:

XML is a flexible, nonproprietary set of standards for tagging information so that it can be transmitted over a network such as the Internet and readily interpreted by disparate computer systems. If implemented broadly with consistent data definitions and structures, XML offers the promise of making it significantly easier for organizations and individuals to (1) identify, integrate, and process information that may initially be widely dispersed among systems and organizations, and (2) conduct transactions based on exchanging and processing such information.[\[GAO\\_1\]](#)

In many respects, the US federal government has embraced XML as a key technology component of its emerging Federal Enterprise Architecture ([FEA](#)). XML and Web services play an important role in the Business Reference Model, the Technical Reference Model and the Data and Information Reference Model of the FEA. XML technology is a central focus of the Emerging Technology group of the Federal CIO Council Architecture and Infrastructure Committee. H.R. 2458, the E-Government Act of 2002, recommends the use of "standards and guidelines for interconnectivity and interoperability"; XML is explicitly called for in the act.

The Office of Management and Budget ([OMB](#)) is actively encouraging agencies to incorporate XML into their E-Government solutions, especially for many of the 25 Presidential E-Government Initiatives. These Initiatives are being encouraged to develop and deposit their XML Schema in a federal registry. When new development or re-development is pursued, XML is expected to be considered for use as the default format for highly structured data as well as for semi-structured information. For legacy repositories that do not directly support XML, legacy to XML mapping and data transformation are called for to foster interoperability.

This paper provides a high-level view of a wide variety of ways in which XML is currently used in the US federal government. In addition to discussing the role of the working groups of the Chief Information Officers Council, the paper covers registry and repository efforts, as well as selected E-Government Initiatives. It highlights some of the many significant efforts underway at various US federal agencies, including (but not limited to): Department of Commerce ([DOC](#)), Department of Defense ([DoD](#)), Department of the Interior ([DOI](#)), Department of Justice ([DOJ](#)), Department of Navy ([DON](#)), Environmental Protection Agency ([EPA](#)), General Services Administration ([GSA](#)), Internal Revenue Service ([IRS](#)), National Archives and Records Administration ([NARA](#)), National Institute of Standards and Technology ([NIST](#)), and Small Business Administration ([SBA](#)).

## Note

Since this paper is exclusively about the US Federal Government's use of XML, any unqualified references to the word "federal" or "government" should be assumed to mean *US Federal Government*. However, the views expressed in this paper are those of the author, as was the selection of topics, and should not be construed to be official positions of the US Federal Government or any of its agencies. Information was gathered from agency Web sites and presentations. Agency representatives or contractors who wish to correct or update details about efforts presented in this paper are encouraged to contact the author directly at <mailto:KSall@SiloSmashers.com> with *XML 2003 Update* in the subject line. Agencies whose XML efforts have not been represented in this paper who wish to be included should likewise contact the author.

## 2. Legislative Motivators for XML

There have been several significant legislative acts since 1996 that have fueled the US federal government's interest in XML technology since 1998. The government's interest in Standard Generalized Markup Language (SGML)<sup>1</sup> This section highlights several acts that suggest the use of XML (directly or indirectly): the Clinger-Cohen Act, the Government Paperwork Reduction Act, Section 508 of the Rehabilitation Act, and the E-Government Act of 2002. More recently, the Emerging Technologies subcommittee of the Architecture and Infrastructure Committee of the CIOC has embraced XML, as evidenced by numerous references in the Federal Enterprise Architecture.

### 2.1. Clinger-Cohen Act

The focus of the Clinger-Cohen Act (CCA), 1996, formerly the Information Technology Management Reform Act, is to streamline Information Technology (IT) acquisitions and emphasize life cycle management of IT as a capital investment. CCA shifts IT procurement authority from GSA back to the agencies and encourages the acquisition of commercial off the shelf (COTS) software and hardware products. CCA also encourages procurement reform, results based management, financial accountability, and business process reengineering. IT investments are to be defined, selected, and managed on the basis of a well-founded business case. This act set the stage for the use of XML in capital IT planning as we will see with the OMB 300 and the FEA. See [CCA].

### 2.2. Government Paperwork Elimination Act

Government Paperwork Elimination Act (GPEA), Title XVII, Public Law 105-277, October 23, 1998, requires federal agencies, by October 23, 2003, to provide for the option of electronic maintenance, submission or disclosure of information, when practicable as a substitute for paper, and use and acceptance of electronic signatures, when practicable. The Act grants full validity to electronic form and encourages a range of electronic signature alternatives, including digital signatures. Implementing electronic transactions and electronic signatures can speed transmission of data and reduce transaction costs. Agencies must consider records management requirements when implementing the GPEA, or whenever they design or augment an electronic information system. For these reasons, agencies including GSA, SBA, EPA, NARA and the Census Bureau are actively pursuing E-Forms (and XForms) technology. GPEA was a natural outgrowth of the Paperwork Reduction Act of 1995, which requires agencies to perform their information resource management activities in an efficient, effective and economical manner. [GPEA]

### 2.3. Section 508 of the Rehabilitation Act

"In 1998, Congress amended the Rehabilitation Act to require federal agencies to make their electronic and information technology accessible to people with disabilities....Section 508 was enacted to eliminate barriers in information technology, to make available new opportunities for people with disabilities, and to encourage development of technologies that will help achieve these goals. The law applies to all federal agencies when they develop, procure, maintain, or use electronic and information technology." Since XML makes it possible and practical to separate data and other content from its presentation, it can be readily used to target devices such as voice browsers that are more accessible for individuals with visual impairments, for example. Both Scalable Vector Graphics (SVG) and XForms were designed with accessibility in mind, so they are especially appropriate technology for meeting Section 508 requirements. These W3C recommendations specifically address accessibility.[Section 508]

### 2.4. H.R. 2458, The E-Government Act of 2002

This legislation establishes a broad framework that requires using Internet-based IT to enhance citizen access to Government information. [EGOV ACT] Among its many provisions, the E-Government Act of 2002:

- Establishes an Administrator of a new Office of Electronic Government (ME) within the Office of Management and Budget, to provide overall leadership and direction to the executive branch on Electronic Government (E-Gov) initiatives.

---

<sup>1</sup>For the early history of SGML and XML in the government, see the appendix (Section 9, "Appendix: Early History of Markup Technologies in the US Federal Government ") by Betty Harvey of ECC.

- Promotes innovative uses of information technology by agencies, particularly initiatives involving multiagency collaboration, through support of pilot projects, research, experimentation and use of innovative technologies.
- Oversees the development of an integrated Internet-based information system by agencies and assist in over-seeing that agency E-Government activities have adequate security.
- Establishes policies to support IT standards, including standards for interconnectivity and interoperability, for categorizing federal electronic information, and for computer system efficiency and security.
- Establishes the (the US Federal) Chief Information Officers Council ([CIOC](#)) to develop recommendations on Government information resources management policies and requirements; share experiences, ideas, best practices and innovative approaches related to information resources management; and assist in identification, development, and coordination of multiagency projects to improve Government performance through use of IT.
- Promote access to high quality Government information and services across multiple channels.

In fact, the E-Government Act *explicitly references XML* twice, as in "maximizing the use of commercial standards as appropriate, including the following: .... standards and guidelines for categorizing federal government electronic information to enable efficient use of technologies, such as through the use of extensible markup language [sic]." The [CIOC](#) established by this act in turn formed the Emerging Technologies subcommittee which has devoted considerable efforts to exploring XML technology via working groups and pilots.

## 3. Federal Enterprise Architecture and the OMB 300

According to an April 2003 report by [SBA](#), the total federal [IT](#) spending for fiscal year 2002 was over \$50 billion. Requested IT budgets for FY2003 and FY2004 were \$57 and \$59 billion, respectively. Clearly, the US Federal Government's investment in IT is Big Business. Methods that can eliminate or reduce redundancies could result in huge savings for taxpayers, or the funds could be applied to other programs that more directly benefit citizens. Such changes are consistent with the [CCA](#).

In April 2002, the [GAO](#) published a report entitled *Electronic Government: Challenges to Effective Adoption of the Extensible Markup Language*. This comprehensive report provided an assessment of the maturity of XML, how XML is being used in government, and recommendations for the next steps in governmental adoption of this technology. In addition, the report encourages federal agencies to explicitly indicate how XML will become part of their enterprise architecture and to build registries and repositories to promote interoperability. See [\[GAO\\_2\]](#).

In June 2000, Mark Forman in [OMB](#) became the first head of federal [IT](#), essentially the Chief Information Officer for the entire US Federal Government. Forman led the effort to define and implement the Presidential E-Government Initiatives [see [Section 5](#), "E-Government Initiatives"] until August 2003. In mid 2002, when Forman was asked which technologies would have an impact on the government over the next year or two, he replied:

XML, without a doubt, because it gives us the ability to collect data once and use it many times.

It allows us to do electronic transactions as opposed to filling out paper forms.<sup>2</sup>

In October 2003, [DOE](#) CIO Karen Evans replaced Forman as OMB's Associate Director for IT and E-Government. Evans is also the vice chair/director of the Federal CIO Council which makes recommendations for federal IT management policy, procedures and standards. Long before becoming the IT Czar, Evans articulated the need for XML in her vision statement to the [CIOC](#) in December 2002:

We must continue developing a governance process for architecture and allow that architecture to drive our investment decisions. We must look at the common transactions within and between government entities and develop standards for those transactions across the government. We

---

<sup>2</sup>The 2002 CNET interview with Mark Forman was a News.com Vision Series Profile and is located at <http://news.com.com/pdf/ne/vs2/forman.pdf>.

will consider publishing a taxonomy for government so we use the same language to describe the same concepts and will develop standards for XML data definitions so the information we create can be shared and accessed easily regardless of its origins.<sup>3</sup>

In July 2003, Evans released a memo with the subject Industry Partnerships which stated:

Instituting an Enterprise Architecture throughout government, and undergoing all of the necessary internal changes that entails, is a daunting task. It is, however, the right thing to do. Making our government more efficient and more effective is what IT is all about. Saving taxpayers dollars through a managed architecture shared by all government elements is in the best interest of our nation....The US government spends more on technology than any other entity in the world. This fact should drive both government and contractors to one inescapable conclusion -- business and government must team together to get results.<sup>4</sup>

This section discusses the Federal Enterprise Architecture (FEA) under development by the OMB and the CIO and the capital IT budget form Exhibit 300 (aka OMB 300, aka OMB Exhibit 300). We will see how the FEA and OMB 300 contribute to reducing redundancies in the IT area and how XML plays a major role in both.

### 3.1. FEA Reference Models

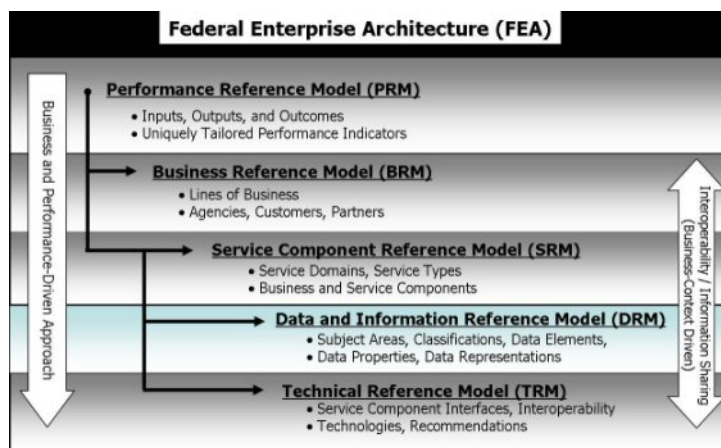
In February 2002, the OMB IT chief, Mark Forman, directed the creation of a Federal Enterprise Architecture Program Management Office (FEAPMO) to begin developing a comprehensive, business-driven blueprint for modernizing the US Federal Government. The emerging Federal Enterprise Architecture (FEA) is a framework that describes the relationship between business functions and the technologies and information that support them. The goal of the FEA is to transform the Federal Government to one that is citizen-centered, results-oriented, and market-based, as well as to maximize technology investments to better achieve mission outcomes. To accomplish this, the FEA expressly encourages collaboration and resource sharing across agencies. It enables the government to identify opportunities to leverage technology and alleviate redundancy, or to highlight where agency overlap limits the value of IT investments. The FEA facilitates horizontal (cross-federal) and vertical (federal, state, and local governments) integration of IT resources, and establishes the "line of sight" contribution of IT to mission and program performance.

The FEA is being defined in terms of five interrelated *Reference Models* designed to facilitate cross-agency analysis and the identification of duplicative investments, gaps, and opportunities for collaboration within and across federal agencies. Major IT investments must be aligned against each reference model within the FEA framework. The reference models required to be used during the FY2005 budget formulation process have been published on the FEAPMO Web site, <http://www.feapmo.gov>. In contrast to many previous federal architecture efforts, the FEA is clearly business-driven. Its foundation is the Business Reference Model, which describes the government's Lines of Business and its services to the citizen independent of the agencies and offices involved. This business-based foundation provides a common framework for improvement in a variety of key areas: budget allocation, horizontal and vertical information sharing, performance measurement, budget and performance integration, cross-agency collaboration, component-based architectures and E-Government.

---

<sup>3</sup>Karen Evans' December 17, 2002 vision statement to the CIO Council is located at [http://www.cio.gov/documents/karen\\_memo\\_12\\_17\\_02.pdf](http://www.cio.gov/documents/karen_memo_12_17_02.pdf).

<sup>4</sup>Karen Evans' July 8, 2003 memo is located at [http://www.cio.gov/documents/industry\\_KE\\_7\\_11\\_03.pdf](http://www.cio.gov/documents/industry_KE_7_11_03.pdf).



XML is frequently mentioned in three of the Reference Models: BRM, TRM and DRM. SOURCE: [www.FEAPMO.gov](http://www.FEAPMO.gov)

### Figure 1. Five Reference Models of the US Federal Enterprise Architecture

The Reference Models of the FEA are described briefly below.

- Performance Reference Model (PRM)** - standardized performance measurement framework designed to characterize performance in a common manner across all federal lines of business. The PRM will help agencies produce enhanced performance information; improve the alignment and better articulate the contribution of inputs, such as technology, to outputs and outcomes; and identify improvement opportunities that span traditional organizational boundaries. Initially released in September 2003.
- Business Reference Model (BRM)** - function-driven framework that describes the Lines of Business, Internal Functions and External Functions performed by the Federal government independent of the agencies that perform them. Major IT investments are mapped to the BRM to promote cross-agency collaboration opportunities. The BRM was the first FEA Reference Model released. The initial version appeared in July 2002 for use with the FY2003 OMB 300; the second version was released in June 2003.
- Service Component Reference Model (SRM)** - common framework and vocabulary for characterizing the IT and business components that comprise an IT investment. The SRM will help agencies assemble IT solutions through the sharing and re-use of business and IT components. In this context, a *component* is a self-contained process, service, or IT capability with pre-determined functionality that may be exposed through a business or technology interface. Initially released in June 2003.
- Data and Information Reference Model (DRM)** - classification system for data to identify duplicative data resources as well as to enable information sharing between agencies. A common data classification model will streamline the processes associated with information exchange both within the federal government and between the government and its external stakeholders. The DRM categorizes the government's information along general content areas specific to BRM sub-functions and decomposes those content areas into greater levels of detail, ultimately to data components that are common to many business processes or activities. Initially released in October 2003.
- Technical Reference Model (TRM)** - framework to describe the standards, specifications, and technologies supporting the delivery, exchange, and construction of business or service components and E-Gov solutions. The TRM unifies existing agency TRMs and E-Gov guidance by providing a foundation to advance the re-use of technology and component services from a government-wide perspective. Initially released in June 2003.

Although the PRM and SRM barely mention XML, the BRM has 11 references to XML; the TRM contains 61 XML references. While many of the references are to specific XML vocabularies, the TRM specifically identifies XML Schema as the primary data format, XSLT as the primary means for data transformation, and XML technology in general for integration, interoperability, and application-to-application interface throughout the FEA to connect disparate systems and information providers. The final version of the DRM was not released at the time of this



writing. However, a draft version contained about 15 XML references. In addition to listing the XML Working Group as one of the practitioners, the draft [DRM](#) leverages ISO 11179, UN/CEFACT, ebXML, Universal Business Language ([UBL](#)), Resource Description Framework ([RDF](#)), OASIS efforts, E-Gov Initiatives, and W3C technology. The draft describes the [DRM](#) as: a framework to support the classification of data and information in respect to how it supports the lines of business and functions within the BRM; a registry that provides multiple levels of granularity to satisfy the re-use of data schemas from multiple stakeholder views; a collection of interrelated, context-driven XML Schemas; a framework that builds upon existing XML Schemas, Data Definition Libraries, and initiatives that exist across the Government (e.g., UN/CEFACT, UBL, ISO 11179, OASIS, current E-Gov Initiatives). The DRM is the only one of the Reference Models not required for FY2004 budget submissions, but it will be for FY2005. See [FEA](#).

## 3.2. OMB 300 as XML Schema

[OMB](#) Circular A-11, Preparation, Submission and Execution of the Budget, must be addressed by each US Federal Government agency every year. This is quite a laborious process because A-11 contains 8 parts and numerous sections, most of which require lengthy and highly detailed responses. This Circular, released to all three branches of the government in July, provides detailed instructions for submitting budget data and materials. Part 7 (section 300) of Circular A-11 establishes policy for planning, budgeting, acquisition and management of Federal capital assets, with detailed instructions on budget justification and reporting requirements for major [IT](#) investments. It is this form that is known as OMB 300 (aka Exhibit 300). Section 53, Information Technology and E-Government, is a companion form that provides reporting requirements for an agency's IT Investment Portfolio. The business case portion of the OMB 300 must include pertinent references to the [FEA](#) reference models, including the Performance Reference Model (PRM). Many of the questions concerning IT relate specifically to the [FEA](#) as well. For example, consider this actual question from the FY2004 OMB 300:

Discuss this major IT investment in relationship to the Technical Reference Model section of the FEA. Identify each Service Area, Service Category, Service Standard, and Service Specification that collectively describes the technology supporting the major IT investment. For detailed guidance regarding the FEA TRM, please refer to <http://www.feapmo.gov>.

Microsoft and other companies have been working with OMB to develop an XML Schema to be used in fiscal year 2005 budget submissions (which will begin to be completed in July 2004). At the time of this writing, the current version of the schema, version 2.92, consisted of 2,280 non-commented lines. The schema for an OMB 300 Capital Asset Plan for IT investment consists of three sections, called `header`, `partOne` and `partTwo` in the XSD. The `partOne` element consists of a sequence of the mandatory children elements `generalQuestions`, `spendingByProjectStages`, `projectDescription`, `justification`, `performanceGoalsAndMeasures`, `programManagement`, `alternativeAnalysis`, `riskInventoryAssessment`, `acquisitionStrategy`, and `projectFundingPlan`. The structure of `partTwo` is a bit more complex, as shown in Figures 2, 3, and 4. The immediate children are `enterpriseArchitecture` (Figure 2 and 3), `securityAndPrivacy` and `part2GPEAQuestions` (both in Figure 4). Figure 2 shows the business and data elements and their nested children.

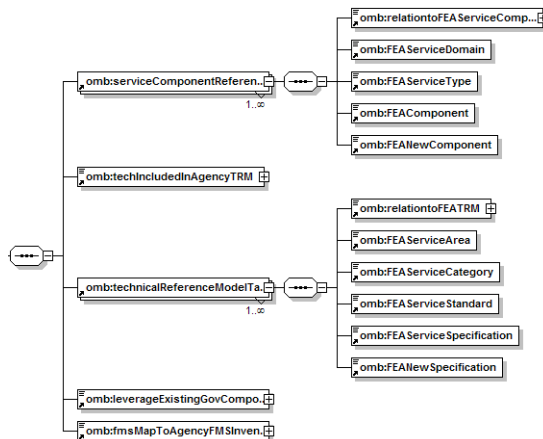


OMB 300 Part Two: business and data Children of enterpriseArchitecture

SOURCE: OMB300v2.92.xsd from www.FEAPMO.gov

### Figure 2. XML Schema Structure of OMB 300, Part Two (A)

Figure 3 shows the children of `applicationComponentsAndTechnology`, namely `serviceComponentReferenceModelTable`, `techIncludedInAgencyTRM`, `technicalReferenceModelTable`, `leverageExistingGovComponents`, and `fmsMapToAgencyFMSInventory`.

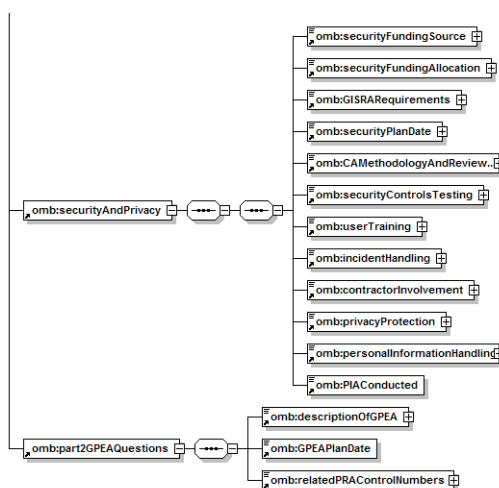


OMB 300 Part Two: Children of applicationComponentsAndTechnology

### Figure 3. XML Schema Structure of OMB 300, Part Two (B)

Figure 4 shows the structure of high level `securityAndPrivacy` and `part2GPEAQuestions` elements.





OMB 300 PartTwo: securityAndPrivacy and part2GPEAQuestions Children

**Figure 4. XML Schema Structure of OMB 300, Part Two (C)**

As you might guess, the OMB 300 schema also contains several enumerations that correspond closely to the Reference Models. For example, there is a `lineOfBusinessString` simple type which is an enumeration of the acceptable three-digit code values from 100 to 410 found in the Reference Models. Similarly, the schema also contains enumerations representing code lists for agencies and bureaus.

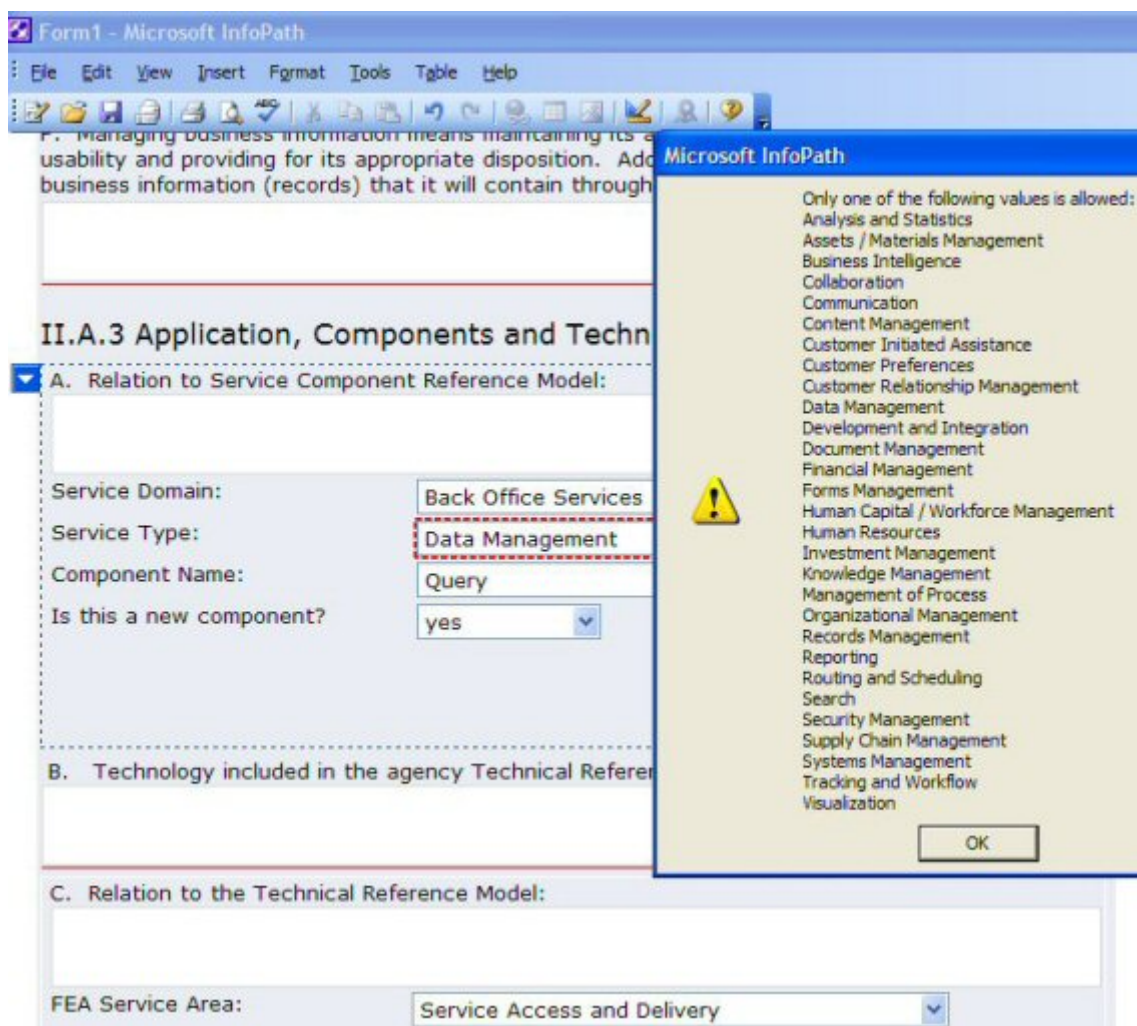
See [\[OMB 300\]](#) for the exhibits and instructions; see [\[FEA\]](#) for the actual XML Schema.

### 3.3. InfoPath Implementation of an OMB 300 E-Form

Microsoft has been working with [OMB](#) to develop a free form that enables generation of the OMB 300 in XML format, specifically by using the XML Schema discussed in the previous section. The Microsoft solution is intended to provide the look and feel of A-11 guidance and can be sent via email to OMB using Microsoft Web services technology.

Their solution is based on the capabilities of InfoPath 2003, a hybrid tool that combines word processor editing with the rigorous data-capture capabilities of forms. InfoPath uses a custom-defined XML Schema (OMB300v2.92.xsd, in this case) to constrain and guide editing the form. InfoPath both consumes and produces XML Schemas and XSLT stylesheets, and is integrated with XML Web services standards such as SOAP. Data can be submitted in XML format via SOAP or by means of the more conventional HTTP POST method. Therefore, the user interacts with a familiar forms interface but the validation logic is driven by an XML Schema and the validated form data is submitted over the Internet.

Figures 5 and 6 are the left and right sides of the OMB 300 InfoPath form. Notice that the PartOne and PartTwo terminology shown on the right of Figure 6 corresponds exactly (except for the initial capital letter) to children elements discussed in the previous section.



SOURCE: Microsoft InfoPath OMB 300 Solution from <http://www.microsoft.com/usa/government/fed/default.asp>

**Figure 5. Microsoft InfoPath OMB 300 E-Form (A)**

In addition to displaying dropdown lists with choices based on enumerations in the XML Schema, additional logic can be added in InfoPath to reflect interdependencies between multiple fields. Error messages can therefore be very precise, as illustrated in Figure 5.



Often there is a combination of experimentation and establishing best practices. In this section, we cover a handful of these activities. Although the guidance documents vary somewhat, topics typically covered include:

- Selecting XML Standards for Project Use
- Importance of International Standards
- Creating ISO 11179 Names
- Creating XML Element Names from Business Terms
- Case Conventions
- Usage of Acronyms and Abbreviations
- Adding Comments and Metadata
- When to Use XML Schema vs. DTDs
- Schema Development Methodology
- When to Use Attributes vs. Elements
- Global vs. Local Elements and Attributes
- Enumeration of Data Values (Code Lists)
- Constraining Data Values
- XML Namespaces
- Web Services Best Practices
- XSLT Best Practices
- Unresolved Issues

For example, one of the most common conventions adopted by most government agencies is to use `UpperCamelCase` for XML element names and `lowerCamelCase` for attributes. Direction to this effect appears in all of the guidance documents discussed in this section.

## 4.1. XML Working Group

The Federal CIO Council's XML Working Group, first chartered in July 2000, is facilitating the efficient and effective use of XML by all government agencies. According to the September 2002 (revised) charter, the XML Working Group is primarily focused on establishing an XML Registry and Repository (for sharing of schema), formulating an XML Developer's Guide including the identification of best practices, fostering "partnerships with key industry and public interest groups developing and implementing XML standards", and encouraging "partnerships among communities of interest/practice involving agencies at all levels of government to capitalize as rapidly and effectively as possible on the potential benefits of XML to citizens and taxpayers".

The XML Working Group, originally co-chaired by Owen Ambur (Department of the Interior, Fish and Wildlife Service) and Marion Royal ([GSA](#), Office of Governmentwide Policy), maintains a very useful XML.gov Web site (<http://XML.gov>) that collects numerous technical presentations, meeting minutes, agency efforts, completed and in-progress documents, maintains an extensive list of XML registries/repositories, and more.<sup>5</sup> Two of the draft documents found at XML.gov are the lengthy *XML Developer's Guide* (April 2002) and the paper entitled *Recommended XML Namespace for Government Organizations* (March 2003)<sup>6</sup>. In this writer's opinion, all government

---

<sup>5</sup>At the time of this writing, there was the possibility that the XML.gov URL might be used for another purpose. If so, the site used by the former XML Working Group would transition to another .gov domain name.

agencies should consider these emerging federal guidelines when they consider how they will implement XML technology. The working group meets on a monthly basis, is open to all agencies and all contractors working with the government, and has a ListServ with roughly 200 members. In October 2003, the CIO Council decided to re-focus its various working groups, so the XML Working Group is now the XML Community of Practice (XML CoP) co-chaired by Owen Ambur and Lee Ellis ([GSA](#), Office of Governmentwide Policy).<sup>6</sup> [[XML WG](#)]

## 4.2. Web Services Working Group

The CIO Council also sponsors a Web Services Working Group which usually shares meeting days and location at the National Science Foundation with the Universal Access Collaboration Workshops (led by Susan Turnbull). The purpose of the Web Services Working Group, formed in December 2002, according to Brand Niemann, the group's Chair, is:

to support the Emerging Technology Subcommittee of the [CIO Council's Architecture and Infrastructure Committee] in its work with the other two subcommittees (Enterprise Architecture Governance and Components) and to produce incubator pilot projects in support of the E-Government Initiatives that use XML Web Services to demonstrate increased accessibility and interoperability.

Niemann has published literally dozens of Activities Reports, Emerging Technology Subcommittee Reports, and CIO Council Reports, in addition to collecting numerous technology and project presentations from working group members on his information-dense Web site, <http://web-services.gov>. Meetings address a very diverse list of topics including ebXML and Universal Discovery, Description and Integration ([UDDI](#)) registries, SOAP and Web Services Description Language ([WSDL](#)), Topic Maps, Semantic Web, VoiceXML, metadata, native XML databases, XForms, E-Forms, [SVG](#), and more. Over a dozen pilot studies, including the E-Forms for E-Gov pilot discussed next, have been sparked by the Web Services Working Group. This group meets on a monthly basis, is open to all agencies and all contractors working with the government (signup by RSVP), and has a ListServ with roughly 200 members. As was the case with the XML Working Group, the Web Services Working Group was recast as the Government Semantic Web Services Community of Practice (SWS-CoP) in October 2003. [[Web Services WG](#)]

## 4.3. E-Forms for E-Gov Pilot

The government is actively investigating *E-Forms* as an offshoot of early efforts of the Web Services Working Group. The E-Forms for E-Gov Pilot was initiated in response to a December 2002 request from the OMB Government-to-Government Portfolio Manager and [FEAPMO](#) through its Solutions Architects Working Group ([SAWG](#)) "to provide information on any electronic forms applications that have been, or are being implemented by federal agencies that meet Section 508 and [GPEA](#) requirements and that reduce or eliminate the problems and redundancies being experienced by the agencies." This was largely motivated by the recognition that over 120 business cases for FY2003 called for E-Forms. (170 E-Forms business cases were submitted for FY2004.) Many vendors are participating in E-Forms efforts, including Microsoft, Adobe, Fenesta, and PureEdge. "Reusable components like E-Forms are at the heart of the US Federal Enterprise Architecture and E-Government and XML standards-based solutions are starting to appear for use across the government."<sup>7</sup>

In addition to supporting [GPEA](#), E-Forms, or at least the W3C XForms, cleanly separates schema (the model), logic, data, and presentation, enabling the same data to be repurposed for different devices, and promoting potential reuse of common form areas (e.g., identification and contact information). Imagine the benefit if all government forms that required contact information used exactly the same model. Now imagine if as a taxpayer or a civil servant you could submit such information once in a secure manner and reuse it whenever you needed to complete a different form! According to [SBA](#) in their *Final Report of the Small Business Paperwork Relief Task Force*, "OMB estimates the cost to provide data required by all approved information collection requests [from citizens and

---

<sup>6</sup>These documents are located at [http://xml.gov/documents/in\\_progress/developersguide.pdf](http://xml.gov/documents/in_progress/developersguide.pdf) and [http://xml.gov/documents/in\\_progress/GS301L1\\_namespace.328.doc](http://xml.gov/documents/in_progress/GS301L1_namespace.328.doc), respectively, although both are expected to be updated by the end of 2003, so the URLs may change.

<sup>7</sup>Brand Niemann, Ph.D., Chair XML Web Services Working Group, U.S. CIO Council. [Web Services: State of the Art in the federal government, July 30, 2003]



businesses] in Fiscal Year 2003 was approximately 8.2 billion hours and \$320 billion." Even a 1% reduction would be significant; a 10% decrease would yield \$32 billion in savings.

OMB contacted Brand Niemann in December 2002 to investigate E-Forms across agencies. Under the leadership of Rick Rogers, CEO of Fenestra, ten subteams were formed in March 2003 to explore various facets of E-Forms: Accessibility, Business Case, Client Specifications, Fixed Content and Behavior, Form Selection, Presentation, Records-Keeping, Schema, Security, and Services. The Selection subteam surveyed a number of government forms before selecting five: SF424 Application for Federal Assistance (GSA), Form 2290 Heavy Highway Vehicle Use Tax Return (Department of Treasury, IRS), SF1012 Travel Voucher (GSA), Form DS-0011 Application for U.S. Passport or Registration (Department of State), and Form BE 10B Benchmark Survey of U.S. Direct Investment Abroad (Department of Commerce). Next, the Schema subteam developed XML Schema to model three of the forms: SF1012, DS-0011, and SF424. Our intention was to uncover XML Schema design and reuse issues. Meanwhile, the Security team wrote an insightful paper highlighting some of the issues surrounding the storage and transmittal of E-Forms data, presentation, and digital signatures. The general security topics were:

- Type of protection needed to meet requirements for authenticity and privacy of electronic records as mandated by federal statute and regulation;
- Practices, processes, and architectures that will ensure the availability and integrity of the E-Forms technology; and
- Security characteristics of specific E-Forms technologies and the security requirements, issues and possibilities of each.

The paper suggests a *security envelope* around E-Forms and describes the functional requirements of E-Forms in terms of authenticity (including data integrity) and privacy. All products of the pilot are available from the Fenestra Web site. [\[E-Forms\]](#)

## 4.4. Department of Navy XML Work Group

The Department of the Navy Chief Information Officer (DON CIO) chartered the [DON](#) XML Work Group to "fully exploit XML as an enabling technology to achieve interoperability in support of maritime information superiority." The primary deliverables being developed by the work group are: a vision document (completed in March 2002); DON XML Developer's Guide; DON XML Policy, Procedures, and Governance Structure; DON XML Implementation Plan; and DON XML Registry/Repository Requirements. Five Action Teams, all overseen by a central Steering Committee are responsible for integrating XML with the DON's command structure and major IT initiatives. Michael Jacobs of the DON CIO is the project lead and Steering Committee chair.

DON released an XML policy statement in December 2002 which generated a fair amount of IT press interest; it was the first comprehensive XML policy from the [DoD](#). The key to interoperability is active involvement in Voluntary Consensus Standards ([VCS](#)) bodies such as W3C, OASIS, ISO, and IETF. In addition to establishing an XML registry, the group has published comprehensive guidance documents. Version 1.0 of the *DON XML Developer's Guide* was released in October 2001, version 1.1 was finalized in May 2002, and version 2.0 was in draft form at the time of this writing. The *DON XML Developer's Guide* is very detailed in the areas of DTD and XML Schema guidance, includes a glossary and eight appendices. This DON document is actually the basis of the CIO Council XML Working Group's *XML Developer's Guide*. In addition to following W3C Recommendations, DON encourages adherence to international standards such as ISO 11179, UBL, and ebXML. See [\[ISO 11179\]](#), [\[UBL\]](#), and [\[CCTS\]](#). A Business Standards Council (BSC) consisting mainly of Functional Namespace Coordinators (FNC) was formed in 2003 to coordinate and harmonize XML components across 23 functional areas. A Technical Standards Council (TSC) is planned for October 2003 to define approved XML technical specifications and to provide technical guidance. [\[DON XDG\]](#)

## 4.5. Environmental Protection Agency XML Design Rules

The [EPA](#) and its state partners are working together to establish the nationwide Environmental Information Exchange Network that will use XML as the primary format for data exchange. Some of the goals of the Exchange Network are to reduce the federal information collection burden on the public and on state and local governments, to promote



data sharing with states and other federal agencies to achieve environmental results, and to effectively use the capabilities of XML technologies in support of the Agency's mission and implementation of EPA's basic programs.

Therefore, in April 2003, EPA's Technical Resource Group of the Exchange Network produced a 200-plus-page document entitled *XML Design Rules and Conventions for the Environmental Exchange Network*, which establishes rules and guidelines for the creation and use of the XML for EPA and its partner states. Topics include: High-Level XML Design Rules, Schema Design Rules, XSL [XSLT] Design Rules, XML Tag Naming Convention (based on ISO 11179 data element naming principles [\[ISO 11179\]](#)), XML Schema Namespace and Versioning Strategy, and XML Enumeration and Code Lists. The authors present various design choices, discuss their relative advantages and disadvantages, then offer recommendations for each issue, both for document-centric and data-centric applications. In many senses, the guide is also a tutorial in that it includes considerable expository material. The EPA guide also contains a handy summary of XML rules and glossary. See [\[EPA XDR\]](#) .

## 5. E-Government Initiatives

In October 2001, 24 Presidential Priority E-Government initiatives were approved by the President's Management Council. In his February 2002 Congressional budget submission, President Bush outlined an E-Government strategy to make it easier for citizens and businesses to interact with the government, eliminate redundant systems, save taxpayer dollars, and streamline citizen-to-government communications. This strategy was one of five key elements of the President's Management Agenda from 2001 to make the Federal government more results-oriented, efficient and citizen-centered. Twenty four (eventually 25) high-payoff, government-wide initiatives were identified to integrate agency operations and [IT](#) investments. To accomplish these objectives, the President urged agencies to concentrate on cross-agency teamwork, rather than individual agency needs. "Effective implementation of E-Government is important in making government more responsive and cost-effective" <sup>8</sup>

The E-Government Initiatives are divided into 4 portfolios and one cross-cutting security initiative. The portfolios and the initiatives they contain are presented below.

- *Government to Citizen (G2C)* - Provide one-stop, on-line access to information and services to individuals.
  - GovBenefits.gov
  - Recreation One-Stop
  - IRS Free File
  - On-Line Access to Loans (E-Loans)
  - USA Services
- *Government to Business (G2B)* - Reduce burdens on business, provide one-stop access to information and enable digital communication using the language of e-business (XML).
  - E-Rulemaking
  - Expanding Electronic Tax Products For Businesses
  - Federal Asset Sales
  - International Trade Process Streamlining
  - Business Gateway
  - Consolidated Health Informatics

---

<sup>8</sup>President George W. Bush, July 10, 2002.

- *Government to Government (G2G)* - Enable federal, state and local governments to more easily work together to better serve citizens within key lines of business.
  - Geospatial One-Stop
  - Disaster Management
  - SAFECOM
  - E-Vital
  - Grants.gov
- *Internal Efficiency and Effectiveness (IEE)* - Modernize internal processes to reduce costs for federal government agency administration.
  - E-Training
  - Recruitment One-Stop
  - Enterprise HR Integration
  - E-Clearance
  - E-Payroll
  - E-Travel
  - Integrated Acquisition Environment
  - E-Records Management
- *E-Authentication* - Reduce the number of credentials by customer segment needed to interact with the Federal government.

In this section, we examine how XML is employed by selected E-Government Initiatives. [\[EGOV INTS\]](#)

## 5.1. Integrated Acquisitions Environment: GSA

With yearly expenditures of roughly \$250 billion on the acquisition of goods and services, the US Federal Government is the world's largest buyer. Although recent procurement reform laws have encouraged federal agencies to undertake various initiatives to streamline acquisitions, the reform has also resulted in duplication of efforts and other significant inefficiencies. Due to the lack of data standards, sharing of business data across agencies is virtually impossible.

The Integrated Acquisition Environment ([IAE](#)) initiative will create a secure business environment that will facilitate and support cost-effective acquisition of goods and services by agencies, while eliminating acquisition inefficiencies. Common acquisition functions that can benefit all agencies, such as the maintenance of information about suppliers (e.g., capabilities, past performance histories) will be managed as a shared service.

The IAE is composed of five modules, one of which is Standard eTransactions. Its objective is to develop a Standard Vocabulary to facilitate exchange of data between and within agencies. In October 2003, the Program Management Office ([PMO](#)) published initial definitions of Standard Information Exchanges and Standard Vocabulary based on an analysis of the existing interfaces for five of the shared IAE systems: CCR (Central Contractor Registration), FPDS-NG (Federal Procurement Data System Next Generation), PPIRS (Past Performance Information Retrieval System), IGT (Intra-governmental Transactions), and FedReg (Federal Registration). A primary goal is to maximize interoperability with other agency, federal lines-of-business, and external (industry, state and local, international) vocabularies.

The process used in defining the IAE Standard Vocabulary is based on the general approach outlined in the FEA DRM: data modeling using Unified Modeling Language (UML), ISO/IEC 11179 data element naming, UBL, and UN/CEFACT Core Component principals. This is followed by XML Schema development to define the precise structure of Information Exchange payloads.

The data modeling and naming process results in Business Information Entities (BIEs), which are data elements with a business context. An Aggregate Business Information Entity (ABIE) is a collection of related pieces of business information that together convey a distinct business meaning in a specific business context. Expressed in modeling terms, it is the representation of an Object Class, in a specific business context. In XML Schema, an ABIE becomes a complex type (e.g., ContactInformation). Ultimately, modular transactional and validation XML Schema will result from combining the ABIEs into Information Exchanges. Back office and agency systems can then apply XSLT stylesheets to map their data elements to or from the IAE Standard Vocabulary. See [IAE].

## 5.2. E-Travel: GSA

The E-Travel Service (eTS), another GSA E-Gov Initiative, will provide a government-wide web-based service that applies world-class travel management practices to consolidate federal travel, minimize cost and produce superior customer satisfaction. The eTS will be commercially hosted to minimize technology costs to the government and guarantee refreshed functionality. From travel planning and authorization to reimbursement, end-to-end service will leverage administrative, financial and information technology best practices to realize significant cost savings and improved employee productivity. The initiative is implementing the FEA vision of common business processes and interoperable data across government agencies.

A significant part of the infrastructure involves establishing a common data model to simplify integration processes associated with travel information exchange. The standardization of data element definitions will greatly improve interoperability among agencies, between agencies and the travel industry, and in defining interfaces for business systems such as financial, payroll, and human resources. To this end, the eTravel PMO conducted an informal data call in February 2003. Nine of the ten representative agencies surveyed responded to the data call, including DOE, DOI, DOJ, Department of Transportation (DOT), EPA, GSA, NASA, USDA, and the Veterans Administration (VA).

Based on the input received, the PMO undertook a data modeling effort that integrated agency responses with eTS RFP Requirements, FTR (Federal Travel Regulation), JFMIP (Joint Financial Management Improvement Program), U.S. Department of State FAM (Foreign Affairs Manual, 6 FAM Volume 100). This data modeling effort confirmed that agencies have very different implementations of these data elements and each agency has additional data elements that the eTravel Service may need to incorporate. A subsequent workshop meeting with representatives from 17 of the 24 BRM agencies further refined the set of data elements from the initial 40 found in the FTR to 385, of which 279 were identified as necessary for interchange with agency business systems. The other 106 elements would likely be relevant for a data warehouse.

Elements identified by the various regulations and by agency subject matter experts were grouped into nine major categories such as traveler profile, travel authorization and planning, reservation and ticketing, and travel vouchers and claims. Each category was further divided into groups. For example, the reservation and ticketing category was divided into four groups: reservations itinerary-transportation, transportation cost, reservations-lodging, and reservations-vehicle. Each group was further subdivided into specific data elements. For example, the data elements in the transportation cost group were: transportation base price, transportation taxes, transportation fee, travel agency fee, transportation cancellation charges, total transportation amount, transport payment method indicator, and transport payment identification number.

All data elements were recorded in a spreadsheet with columns for typical data dictionary information such as description, data type, optional or required or dependent, cardinality, and constraints. The complete set of 385 data elements was provided as GFI (Government Furnished Information) to offerors bidding on the RFP in May 2003. This was accompanied by instructions indicating that standard data elements for eTS data exchange should adhere to the names and characteristics identified in the spreadsheet. Data element naming should follow the guidelines of the CIO Council's XML Working Group draft XML Developer's Guide, which strongly recommends the use of either business terms or names that comply with ISO/IEC 11179. The groups identified in the spreadsheet correspond in general to complex types of an XML Schema. The intention was for offers to create XML Schema for

validation and transactions, in accordance with "standard data output" and "application integration" performance objectives of the RFP to simplify agency integration with eTS. See [\[ETravel Service\]](#).

## 5.3. Business Gateway: SBA

The Business Gateway (formerly known as Business Compliance One-Stop, BCOS) Initiative, managed by [SBA](#), is shifting its focus to E-Forms and [GPEA](#) requirements. With 17 billion responses to federal forms per year, forms account for half of the government's \$320 billion total annual paperwork burden on citizens and business. Earlier BCOS efforts were more focused on reducing the burden on businesses by making it easy to find, understand, and comply with relevant laws and regulations at all levels of government. In both cases, the intention is to save businesses billions of dollars due to the reduced time investment in government-to-business (G2B) interaction. Near future plans call for the creation of a single Business Gateway portal into the Federal cross-agency portal for businesses, integrating the content and functionality of [SBA.gov](#) (<http://SBA.gov>), [BusinessLaw.gov](#) (<http://BusinessLaw.gov>), the U.S. Business Advisor (<http://www.business.gov>), and [FedForms](#) (<http://www.fedforms.gov>) into one comprehensive site called [Business.gov](#) (<http://Business.gov>).

SBA in collaboration with GSA will also create an "E-Forms gateway" for federal forms systems. The initiative plans to develop XML Schemas to streamline, harmonize, and automate information collection requirements that affect industry verticals (i.e., food, chemicals, and health care). Forms development will include an approval process and form models will be contributed to an XML Schema repository to ensure data compatibility and reuse. The solution architecture calls for a government-wide E-Forms framework that will also directly support [GPEA](#) requirements. Savings will be realized by reducing the time to complete forms, by collecting data once and reusing it repeatedly, by streamlining decision-making processes, and by providing quicker responses to questions and service requests. XML will be the syntax for data exchange. Forms and data will be routed to agencies in XML format, possibly via Web services. By saving half of all users just 5 minutes each, the reduction in filing burden on citizens and businesses would total \$28 billion and greater government efficiency by improved data quality and electronic data collection and dissemination would save \$56 billion, according to SBA. The long-term goal of the Business Gateway is to become the "one-stop" services portal for G2B interaction for the federal government. See [\[Business Gateway\]](#).

## 5.4. E-Records Management: NARA

[NARA](#) is the managing partner of the E-Government Electronic Records Management Initiative. The main objective of the initiative is to provide guidance concerning tools and standards for electronic records management ([ERM](#)) applicable government-wide. This will enable agencies to transfer electronic records to [NARA](#) in a variety of data types and formats so that they may be preserved for future use by the government and citizens. Goals include establishing practices to assure the integrity of e-records and information; employing [ERM](#) to support interoperability, timely and effective decision making, and improved services to customers; and to provide the tools for agencies to access e-records for as long as required and to transfer permanent e-records to [NARA](#) for preservation. Department of Energy ([DOE](#)), [EPA](#), and [DoD](#) are actively engaged in projects of this initiative.

The Records Management Metadata and Schema Project of this initiative is focused on providing tools to agencies for the transfer of records to [NARA](#). Its main objectives are to identify the metadata the agencies need to transfer electronic records using XML and to create the XML Schema to encapsulate those metadata element. In August 2003, initial XML Schemas capable of supporting automated transfer and accessioning of electronic records were deposited in the GSA-NIST Proof-of-Concept XML Registry/Repository, discussed in Section 6.1. See [\[NARA\]](#).

## 5.5. Recreation One-Stop: DOI

Recreation One-Stop, managed by the Department of the Interior (DOI), provides a single-point of access, user-friendly, web-based resource to citizens, offering information and access to government recreational sites. County and state data were added to [Recreation.gov](#) as part of the Government Without Boundaries (GWOB) initiative started in September 2002. Eventually data for over 2,800 recreation sites was added. In summer 2003, [Recreation.gov](#) launched with enhanced user interface and mapping capabilities. The GWOB Schema Repository for Parks and Recreation schema is located at <http://www.gwob.gov/parks/PRAAppSchemas.htm>.

This initiative is developing the Recreation Markup Language ([RecML](#)) data standard to improve data exchange among federal, state, tribal, local, and non-government organizations. RecML is a voluntary standard that will be adopted by consensus reached by government agencies and non-government organizations interested in recreation-related travel and tourism. The data standard from September 2003 defines terms for recreation areas (parks), facilities (trails, campgrounds, etc.), activities (hiking, wildlife viewing, etc.), alerts (temporary closures) and transactions (reservations, fees, etc.). Eventually RecML will be integrated with other data standards for listing events, responding to customer inquiries, and fulfilling orders for maps and publications so computer systems can exchange data easily via the Internet (i.e., with Web services). RecML will streamline the processes required to update websites and print new editions of recreation-related publications. Federal agencies already use RecML to disseminate their recreation data that has been compiled in the Recreation.gov database. Through Recreation One-Stop and RecML, citizens will be able to:

- Obtain information about parks, museums, monuments, historical landmarks, and other recreation sites including hours of operation, fees, public accommodations and services.
- Make reservations, order passes, and conduct other service transactions on-line (e.g., through FirstGov.gov).
- Access government-collected data relevant to recreation activities.
- Link to related information and services provided by non-governmental partners.

See [\[Recreation One-Stop\]](#).

## 6. Registry, Repositories and Web Services

This section highlights a few of the many existing US Federal Government efforts to develop registries and repositories for the sharing of XML elements and XML Schema. Although there is presently no single, government-wide logical registry, that is the direction the government appears to be headed. For a comprehensive list of XML registries and repositories, see [\[XML Registries\]](#).

The full benefits of XML will be achieved only if organizations use the same data element definitions and those definitions are available for partners to discover and retrieve. A registry/repository is a means to discover and retrieve documents, templates, and software (i.e., objects and resources) over the Internet. The registry is used to discover the object. It provides information about the object, including its location. A repository is where the object resides for retrieval by users.<sup>9</sup>

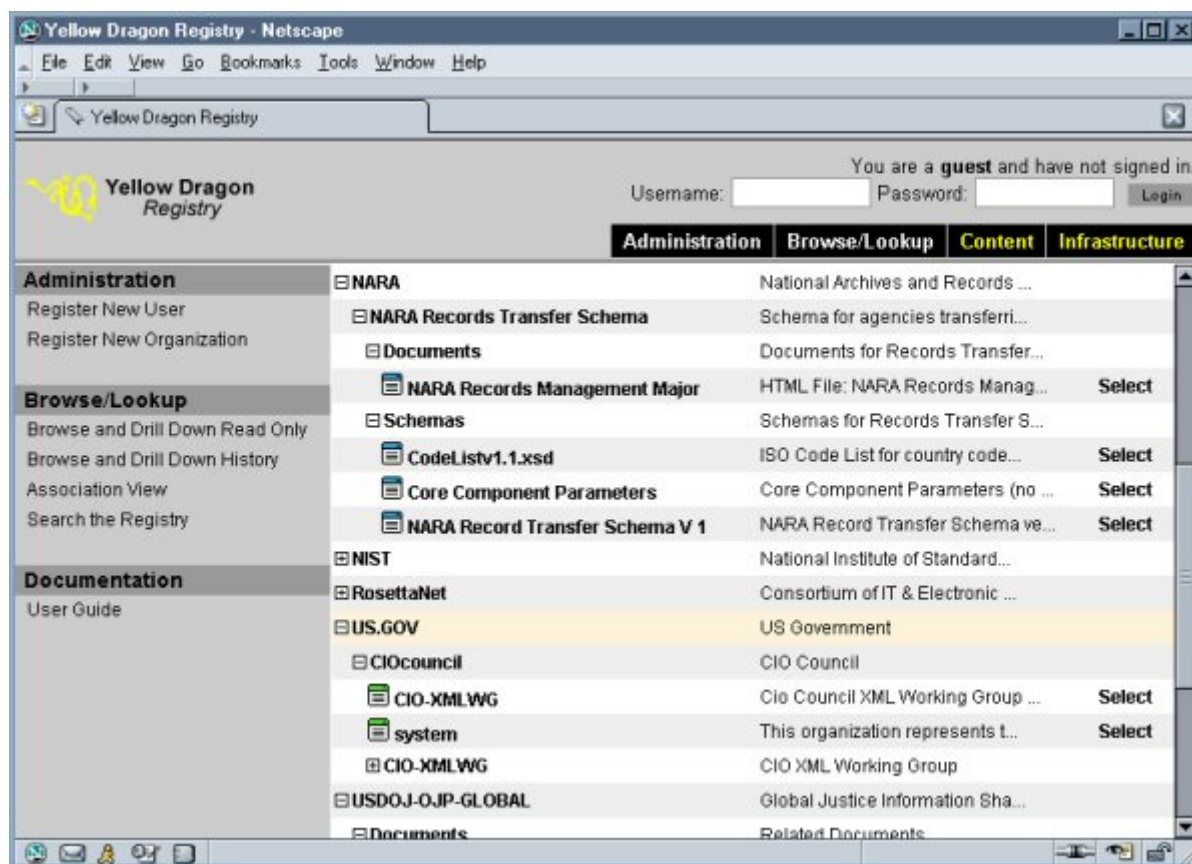
### 6.1. GSA-NIST Proof-of-Concept XML Registry/Repository

The Proof-of-Concept XML Registry/Repository is the third incarnation of a federal registry pilot co-sponsored by [GSA](#) and [NIST](#). In 2002, GSA contracted Booz Allen Hamilton to compile a detailed capital asset plan and business case (see <http://xml.gov/documents/completed/bah/registryBusinessCase.htm>). At the time of this writing, a few XML Schemas, DTDs, code lists, and supporting documents have been registered by [DOJ](#), [NARA](#), [NIST](#), and the XML Working Group. The ebXML registry software shown in Figure 7 is from Yellow Dragon, a spinoff of GoXML. [\[NIST XML Registry\]](#)

---

<sup>9</sup>Owen Ambur, Co-Chair XML Working Group, U.S. CIO Council [XML.gov Registry and Repository page, <http://xml.gov/registries.asp>.]





Drill-down in of Yellow Dragon Registry in Association View. SOURCE: <http://xmlregistry.nist.gov:8080/index.jsp>

**Figure 7. GSA-NIST Proof-of-Concept XML Registry/Repository**

## 6.2. GSA Consolidated Component Repository

At the time of this writing (late September 2003), the [GSA](#) Office of Electronic Government (ME) announced an Request For Proposal ([RFP](#)) for the creation of a Consolidated Component Repository ([CCR](#)). The [CCR](#) will be a registry and repository, intended to evolve into a collaborative environment that will facilitate the creation and sharing of reusable components in the [FEA](#). Such components include, but are not limited to: E-Government Initiatives, data models, XML artifacts (e.g., XML Schema and XSLT stylesheets), COTS product configuration files, software executables, software source code, scripts, and supporting documentation.

A related [RFP](#) was announced for the Business Gateway E-Forms Project, co-sponsored by GSA and the Small Business Administration, intended to provide a simplified gateway for communication with small business. It will include a platform for registration of XML-enabled electronic forms, possibly using a Tamino native XML database. The Business Gateway is viewed as an opportunity for agencies to fulfill their [GPEA](#) requirements.

## 6.3. EPA's Environmental Data Registry

The Environmental Data Registry ([EDR](#)) is a comprehensive reference for the definition, source, and uses of environmental data. The registry supports the creation and implementation of data standards that are designed to promote the efficient sharing of environmental information among [EPA](#), states, tribes, and other information trading partners. The registry does not contain environmental data; instead it provides descriptive metadata for interpretation of the data. Finalized data elements are categorized as Biological Taxonomy, Chemical Identification, Date, Enforcement/Compliance, Facility Identification, Latitude/Longitude, Permitting, SIC/NAICS, and Tribal Identifier. Data standards under development include Contact Information, Environmental Laboratory Results, Federal Facility Identification, Permitting Information, and Reporting Water Quality Results for Chemical and Microbiological Analytes. The related Environmental Information Exchange Network has initiated an interim



XML Registry for sharing information about XML Data Exchange Template (DETs), XML Schemas, namespaces, Web Services Description Language (WSDL) files, and other supporting files needed to map data flows between partners. Registration of new schema information in the registry is accomplished using a Microsoft Excel spreadsheet data template. The registry supports a variety of searching methods. [\[EPA Registries\]](#)

## 6.4. DoD XML Registry

The [DoD](#) has had its own XML and Metadata Registries in one form or another since late 1998. Defense Information Systems Agency ([DISA](#)) is responsible for data-related infrastructures that promote interoperability and software reuse in the secure, reliable, and networked environment planned for the DoD's Global Information Grid (GIG). The registries are part of the DoD Common Operating Environment (COE). There are several related parts:

- *DoD Metadata Registry and Clearinghouse* - Developers can access registered XML data and metadata components, COE database segments, and reference data tables and related metadata information (e.g., Country Code and US State Code). This increases the DoD's core capabilities by integrating common data, packaging database servers, implementing transformation media and using Enterprise data services built from "plug-and-play" components and data access components.
- *DoD XML Registry* - The XML registry is a vital component in the implementation of shared data exchanges. This Registry enables the consistent use of XML, both vertically within projects and horizontally across organizations. The DoD XML Registry constitutes guidance in the generation and use of XML among DoD communities of interest and is the authoritative source for registered XML data and metadata components.
- *Data Element Registry* - This registry constitutes guidance in the generation and use of XML among DoD communities of interest. It is the authoritative source for registered XML data and metadata components. The registry is searchable and may be downloaded as an MS Access database (updated monthly).
- *Reference Data Set Gallery* - Data sets are used across the DoD as a uniform representation of data. A data set is contained in an ASCII delimited file and is associated with a functional steward or authoritative source. Examples are Country Code, US State Code, Purchase Order Type Code, and Security Classification Code.

Due to DoD Operations Security changes, a perspective user must first establish a government account or government-sponsored account to access the majority of the features of the Metadata Registry (which includes the XML Registry). After registration, a user may browse, search, substring search, and retrieve data. Figure 8 shows the search interface, in this case looking for an XML Schema datatype named longitude in all available XML Schema ([XSD](#)) files. The search results page (not shown) leads to a details page like the one shown in Figure 9 which, in this case, indicates the range of values and the base data type. The searcher can easily register for automatic email notification when the document is updated or removed.

DoD XML Registry - Search the Registry - Netscape

File Edit View Go Bookmarks Tools Window Help

DoD METADATA REGISTRY AND CLEARINGHOUSE

DoD XML Registry

Session validated for Kenneth Sall [Logout]

Search the Registry ?

Word or Phrase: longitude

Case-Sensitive? ☐ (Check for YES)

Search On: Name Field

Namespace: ☒ Search All Namespaces  
☐ Search One or More of the Selected Namespaces

DUNWPL - Department of Navy Weapons Planning & Control (Sub-Namespace)  
DONS&T - Department of Navy Scientific & Technology (Sub-Namespace)  
EB - eBusiness  
ENT - DoD Enterprise  
FIN - Finance and Accounting  
GEO - Geospatial and Imagery  
XML - General XML Schema Data Resources

Info Resource Type: XML Schema DataType

Status: All

Version: All

Submitter: All

Start Date: End Date

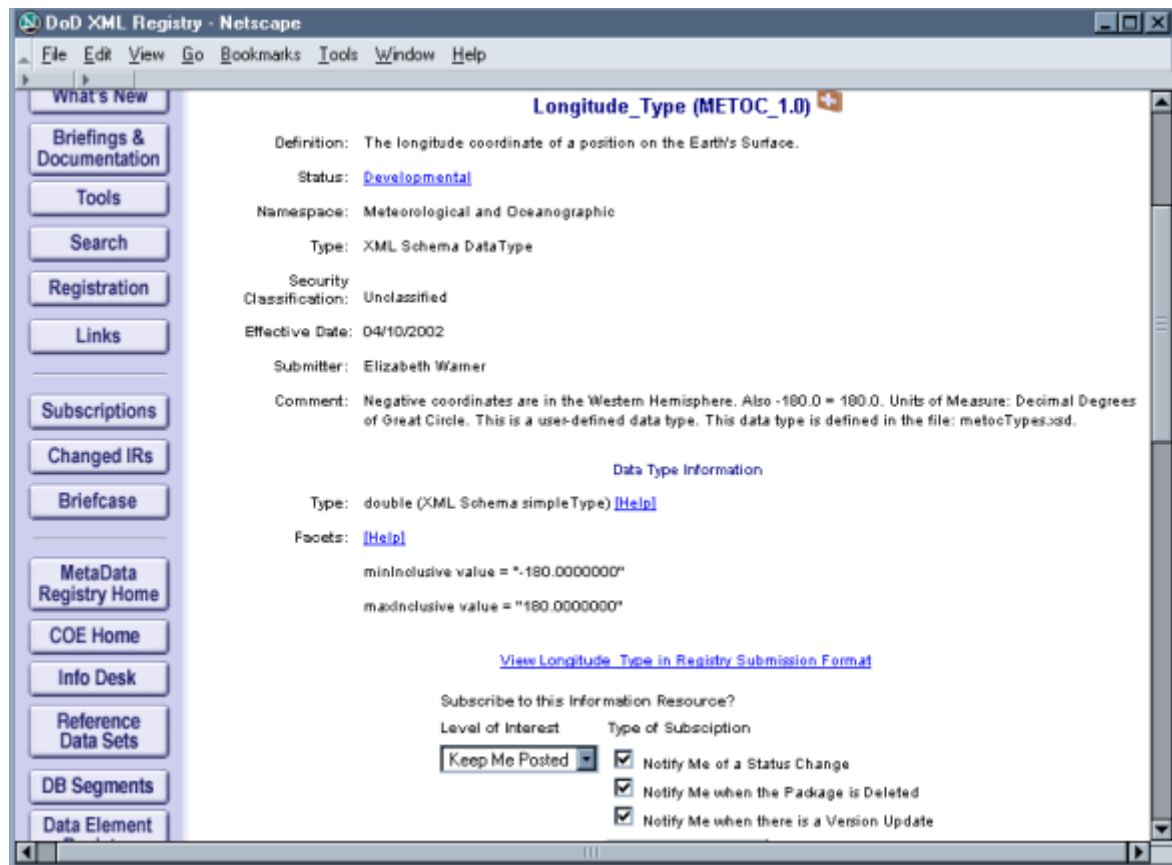
Effective Date: 11/17/1998 10/15/2003

Search

SOURCE: <http://diides.ncr.disa.mil/xmlreg/user/index.cfm>

**Figure 8. DoD XML Registry Search**

The DoD Registry introduces the concepts of Information Resource and Namespace. An *Information Resource* is the generic term used for each object registered in the XML Registry, such as an XML element, attribute, simple type, complex type, DTD, XSD, XSLT stylesheet, etc. A *Namespace* represents a collection of data constructs that share a common context within a Community of Interest (COI), a collection of people, agencies, activities, and system builders who share an interest in a particular domain or practical application and need common data representations. For the DoD Registry, Namespaces are more than just a means for avoiding element name conflicts. There is also a definite administrative role; DoD requires a Point of Contact who functions as Namespace Manager, responsible for assigning status levels to each of the Information Resources as deprecated, developmental, operational, or retired.



SOURCE: <http://diides.ncr.disa.mil/xmlreg/user/index.cfm>

**Figure 9. Drill-down to Longitude Definition from Search Results**

Developers are encouraged to search through the registered Information Resources, adopt them whenever possible, and subscribe to notifications. If the existing Information Resources do not meet the needs of the developer, he can submit a proposed Information Resource to a Namespace manager. Registering Information Resources requires the creation of a submission package (ZIP file) containing the documents, elements, schema, etc. to be registered. The package must contain a file called Manifest.xml, which conforms to registry.xsd or registry.dtd. Manifest.xml file defines all the new Information Resources, as well as the associations with existing Information Resources within the XML Registry. All new Information Resources defined in the submission package should belong to the same Namespace, although existing Information Resources from other Namespaces may be reused. See [\[DoD XML and Metadata Registries\]](#)

## 6.5. DISA Registry Initiative

DoD Data Interchange Standards Association (DISA) supports EDI and XML transactions. (Note that DISA is different from the DISA in the previous section ;-). DISA Registry Initiative (DRIVE) is a registry effort based on the ebXML specifications for business items. The DRIVE registry will include business process models, XML Schemas, DTDs, and industry-specific code lists. DRIVE will also index ebXML-compliant company profiles to interconnect potential trading partners. The intention is to incorporate SOAP messages, UN/CEFACT Core Components and other ebXML developing standards. Since more and more registries are supporting the ebXML Registry specifications, interoperability with these registries is assured. EbXML is an international effort by OASIS and UN/CEFACT to develop common specifications to conduct business with any other company in any industry. It has been endorsed by some of DISA's affiliated organizations, including Accredited Standards Committee (ASC) X12 (EDI focused) and Open Travel Alliance (OTA). See [\[DRIVE\]](#) and [\[CCTS\]](#).

## 6.6. Web Services and Registries Pilot

A number of companies are collaborating on a Web Services and Registries pilot under the Web Services Working Group. The pilot's primary objective is to demonstrate interoperability of [UDDI](#) and ebXML registries for Web services-based collaborations. The secondary objective is to demonstrate specific capabilities of UDDI and ebXML Registry for Web service description registration, maintenance and discovery. Participation from vendors whose products implement UDDI Version 3.0 and OASIS/ebXML Registry Version 2.5 was encouraged starting in July 2003.

The pilot will demonstrate the capability of each registry type to *reach through* to the other registry type to access an object, such as a [WSDL](#) document, that resides in the other type of registry. Use cases will involve at least one each of ebXML registry-enabled and UDDI-enabled trading partners. Each registry will register and maintain WSDL documents. A UDDI-enabled trading partner will reach through to access a Web services description that is maintained in an ebXML registry by registering a record of the WSDL in their UDDI registry and reaching through to the ebXML registry to access the record on demand. Plans for the pilot outline a number of use cases. At the time of this writing, approximately ten vendors were participating in a collaborative demonstration. A public, live demonstration is planned for early 2004. See [\[Web Services and Registries Pilot\]](#).

## 7. Selected Specialized Applications

### 7.1. IRS Publications, Instructions, and Guidance in XML

[IRS](#), a long time proponent of SGML, has also embraced XML. A number of XML efforts are reaching maturity at IRS, possibly by the end of 2003. These include an agency-wide XML registry with schemas; the 1040 family of instructions in XML; compliance with Section 508 requirements; XML versions of taxpayer guidance documents; and a common agency-wide XML document format.

In addition to developing XML Schema for data interchange, the IRS has produced the entire 1040 set of instructions as XML. Epic Editor from Arbortext was used to author the documents, transformations were accomplished using XSLT and Omnimark, and PDF versions were produced using Datalogics Pager. This accomplishment was achieved despite both the complexity of the documents and the need to continue to produce paper versions of the instructions.

To comply with Section 508 accessibility requirements, IRS employed GH Braille, Inc. to author textual descriptions of more than a thousand graphics. Prior to that effort, over 2,000 graphics had been converted to tables. IRS customized its authoring environment and modified its review process to gain acceptance of the new accessibility requirements from its authors. As a result, all tax form instructions and information publications (roughly 475 documents) will have XML-based graphic descriptions and table summaries maintained by the authors of the tax products.

The [IRS](#) Office of Chief Counsel has recently launched the IRS's newest XML application: beginning in October 2003, all taxpayer guidance documents, including revenue rulings, procedures, announcements, and guidance letters will be published as XML. By 2005, over 50,000 documents from that office will be transformed behind the scenes from a customized Word format to XML. The IRS also converted its weekly policy magazine, the Internal Revenue Bulletin, to a pure XML application. This document, which pronounces the official IRS position on a wide range of tax issues, averages from 40 to 200 pages per week and is being produced and processed by Arbortext products. The majority of IRS public documents will be transformed from their current DTD to DocBook 4.2 using XSLT for a standard conversion and display as HTML on <http://www.irs.gov/>. See [\[IRS\]](#) and [\[Section 508\]](#).

### 7.2. House of Representatives XML and Legislative Documents

For several years, government employees of the U.S. House of Representatives have been developing a set of DTDs and, more recently, XML Schemas to model bills, resolutions, and House members. According to the Web site at <http://xml.house.gov>, the Secretary of the Senate and the Clerk of the House have worked together with the Library of Congress and the Government Printing Office to create DTDs for use in the creation of legislative documents using XML, under the direction of the Senate Committee on Rules and Administration and the House Committee on Administration. Although designated as drafts, these models are fairly complex with over 100 elements

and data dictionary. These documents are in the public domain. The paper *Drafting Legislation Using XML at the U.S. House of Representatives* at <http://xml.house.gov/drafting.htm> discusses the history of drafting legislation at the House and current efforts to create an authoring application. Joe Carmel, the author of the paper, states:

Beginning in 1997, the House and the Senate, along with the other Legislative Branch agencies, began an investigation of the use of SGML and later XML as a data standard for the exchange of legislative documents. By December 2000, the Committee on House Administration and Senate's Rules Committee adopted XML as the primary standard for the exchange of legislative documents between the House, Senate, and other legislative branch agencies....[T]he House has been using the XML authoring environment for House-only resolutions since January 2001 and began drafting bills in XML in September 2002. The House plans to draft over 95% of introduced bills in XML by January 2003.

### 7.3. Justice XML Data Dictionary (JXDD) and Justice XML Registry/Repository (JXRR)

Justice XML Data Dictionary ([JXDD](#)) is a set of approximately two thousand unique, standardized data elements needed by Justice Information Systems at the state and local levels. [JXDD](#) is the result of ongoing efforts of the Office of Justice Programs (OJP) and the Global Justice Information Network Advisory Committee (Global), with support from the U.S. Department of Commerce's Telecommunications and Information Administration (NTIA) and the Georgia Tech Research Institute (GTRI).

An object-oriented version of the dictionary, [JXDD](#) version 3.0, was developed by GTRI. They accumulated approximately 16,000 data elements from 35 different Justice and Public Safety sources. By removing the redundancies and resolving semantic differences, the set was reduced to 2,000 unique data elements which were then used to define approximately 300 core data object types (reusable components). Inputs to the data model include the Reconciliation Data Dictionary (RDD), RISSIntel schema, Rap Sheet schema, CourtFiling DTD, Driver History schema, Incident Report schema, Justice Info Exchange Model (JIEM) data sets, Arrest Warrant schema, Charging Document schema, Sentencing Order schema and others. These components have inherent qualities enabling access from multiple sources and reuse in multiple applications. Standardization of the core components creates the potential for increased interoperability across justice and public safety information systems.

The ability to represent objects and their relationships is key to creating a fully beneficial justice information sharing tool. For example, a person can be described in terms of physical characteristics, biometric data, or sociocultural descriptors. When a person is considered in the context of another object, such as a vehicle, a variety of relationships can be expressed. Does this person own the vehicle or is he the driver? Was he hit by the vehicle or did he steal it? [JXDD](#) uses RDF to help express object relationships.

GTRI employed a Microsoft Access database with various export capabilities that simplify distribution of the data dictionary to the [JXDD](#) development team. However, to make effective use of [JXDD](#), Global has proposed the creation of Justice XML Registry/Repository ([JXRR](#)) to house the Data Dictionary, various XML artifacts (e.g., XML Schemas, DTDs, XML Instance Documents, stylesheets), metadata, related business processes, and other supporting components. In addition to providing access to technical tools, the JXRR will facilitate true information sharing among Justice and Public Safety practitioners far beyond capabilities of earlier techniques.

The [JXDD](#) is based on many US government and international standards, including:

- XML.GOV Draft Federal XML Developer's Guide [[XML WG](#)]
- Dept of Navy Draft XML Registry Requirements
- DoD 5015.2-STD Design Criteria Std for E-RMS Apps
- FBI Electronic Fingerprint Transmission Spec v7
- ANSI / NIST Data Format for Interchange of Fingerprint, Facial, and SMT
- ISO / IEC 11179 Specification and Standardization of Data Elements [[ISO 11179](#)]

- UN / CEFACCT ebXML Core Components Technical Spec 1.85 [CCTS]
- OASIS XML Common Biometrics Format Committee
- W3C XML Schema Specification
- W3C RDF and RDF Schema Specification

The Web site provides a wealth of information about the data model, the numerous schemas and code lists, the development process, and background information about partner organizations. [JXDD]

## 7.4. Generalized Instrument Design System (GIDS): Census Bureau

The U.S. Department of Commerce/Census Bureau looked to Fenestra (<http://www.fenestra.com>) to explore reuse capabilities and survey form redesign for the 2002 Economic Census initiative, which takes place every five years. Fenestra's solution was based on separation of content and presentation using XML, XSL, and XSLT with content metadata stored in an Oracle database and maintained by a standard Oracle Developer client/server application. Managing the content separately from the layout enabled Fenestra to produce approximately 80% of the survey layouts automatically via rules and style sheets. This automatic formatting process is implemented by a server application called Generalized Instrument Design System (GIDS) AutoFormatter. For the remaining 20% of survey questions that didn't follow regular rules, Fenestra developed a content-aware layout designer application called GIDS Forms Designer. The GIDS software has been deployed to hundreds of workstations at Census, and is being used to produce both paper and electronic versions from the same content for over 650 surveys distributed to over six million businesses. These techniques resulted in a reduction from 10,000 pages to 200 pages. GIDS also uses an ISO 11179 Metadata Registry to design forms and an import wizard for cross-mapping data elements. See [GIDS].

## 7.5. EPA Emergency Response VoiceXML Application for DHS

In 2002, EPA undertook several successful projects using VoiceXML including a Digital Talking Book on CD-ROM and VoiceXML over the Web to access the Local Emergency Planning Committee (LEPC) Database at <http://www.epa.gov/ceppo/lepclist.htm>. LEPC provides

a forum for emergency management agencies, responders, industry and the public to work together to understand chemical hazards in the community, develop emergency plans in case of an accidental release, and always look for ways to prevent chemical accidents. Local industries must provide information to LEPCs about chemical hazards, LEPCs are required by law to make this information available to any citizen who requests it.<sup>10</sup>

VoiceXML is typically delivered to a voice browser with audio output (computer-synthesized or recorded) and audio input (voice and/or keypad tones). Telephones can act as the interface. EPA's Chemical Emergency Preparedness and Prevention Office (CEPPO) maintains a FileMaker Pro contact database with over 3000 LEPC. This database can be queried by zipcode to locate LEPCs in the given area. To obtain LEPCs in zipcode 20794, for example, use the URL [http://130.11.53.73/lepc/FMPro?-db=LEPC.FP5&-format=fmp\\_xml&zip\\_lepc::zip\\_code=20794&-find=](http://130.11.53.73/lepc/FMPro?-db=LEPC.FP5&-format=fmp_xml&zip_lepc::zip_code=20794&-find=). This query returns a VoiceXML document. You can also access the telephone interface by dialing 1-800-303-9987 and saying or keying in the application ID 713589. The response consists of voice output indicating nearby LEPCs which should match the XML output via URL access.

In March 2003, Martin Smith announced that he was coordinating an effort under the CIO Office of Department of Homeland Security (DHS) to encourage the effective use of Web services and related distributed-computing technology within DHS. In May 2003, EPA LEPC VoiceXML Web Service, EPA Facility Data XML and VoiceXML

---

<sup>10</sup>As documented at <http://xml.gov/documents/completed/VoiceXMLaward.htm>. US Environmental Protection Agency, Natural Language Interface to Web Content (VoiceXML and the FedGov XML Content Network) was recognized for technical excellence by the CIO Council in March 2002.

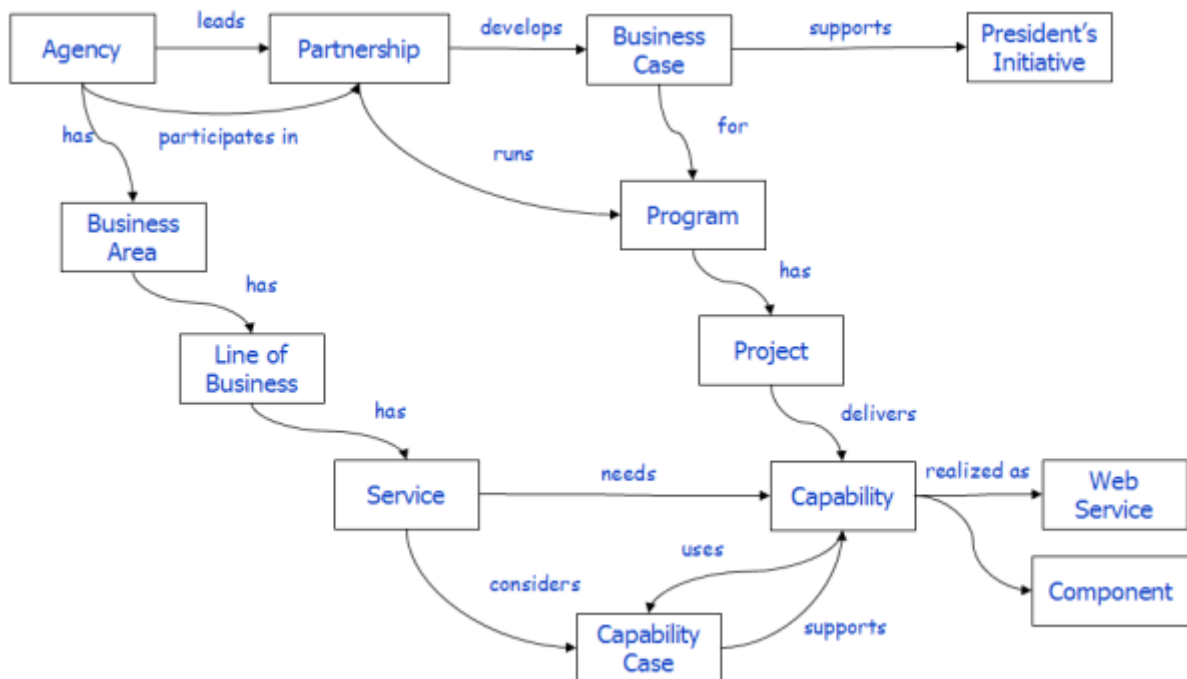


Web Services, and the MetaMatrix System for Model-driven Integration with Enterprise Metadata Pilots were "graduated" to the Disaster Management E-Gov Initiative and the [DHS](#) CIO Office. [\[VoiceXML\]](#)

## 7.6. Semantic Web

In September 2003, a *Semantic Technologies for E-Gov Conference* was held at the White House Conference Center. The event, organized by Brand Niemann and Ralph Hodgson of TopQuadrant drew over 130 attendees from the Army, Census Bureau, CIA, DIA, [DOE](#), [EPA](#), [GSA](#), [IRS](#), Navy, [NARA](#), National Aeronautics and Space Administration ([NASA](#)), National Security Agency ([NSA](#)), National Science Foundation ([NSF](#)), Social Security Administration ([SSA](#)), US Department of Agriculture ([USDA](#)) and [USPTO](#), as well as from many government contractors. See [http://www.topquadrant.com/conferences/tq\\_proceedings.htm](http://www.topquadrant.com/conferences/tq_proceedings.htm) for several valuable presentations and white papers. TopQuadrant also prepared a poster based on one of the white papers, entitled *Delivering Value-Driven Semantic Solutions for Government*; see [http://www.topquadrant.com/conferences/tq\\_poster.htm](http://www.topquadrant.com/conferences/tq_poster.htm). Michael Daconta, coauthor of a book on the Semantic Web, presented his "Declaration of Data Independence". [See Daconta, Obrst, and Smith, *The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management*, Wiley Technology Publishing, June 2003.]

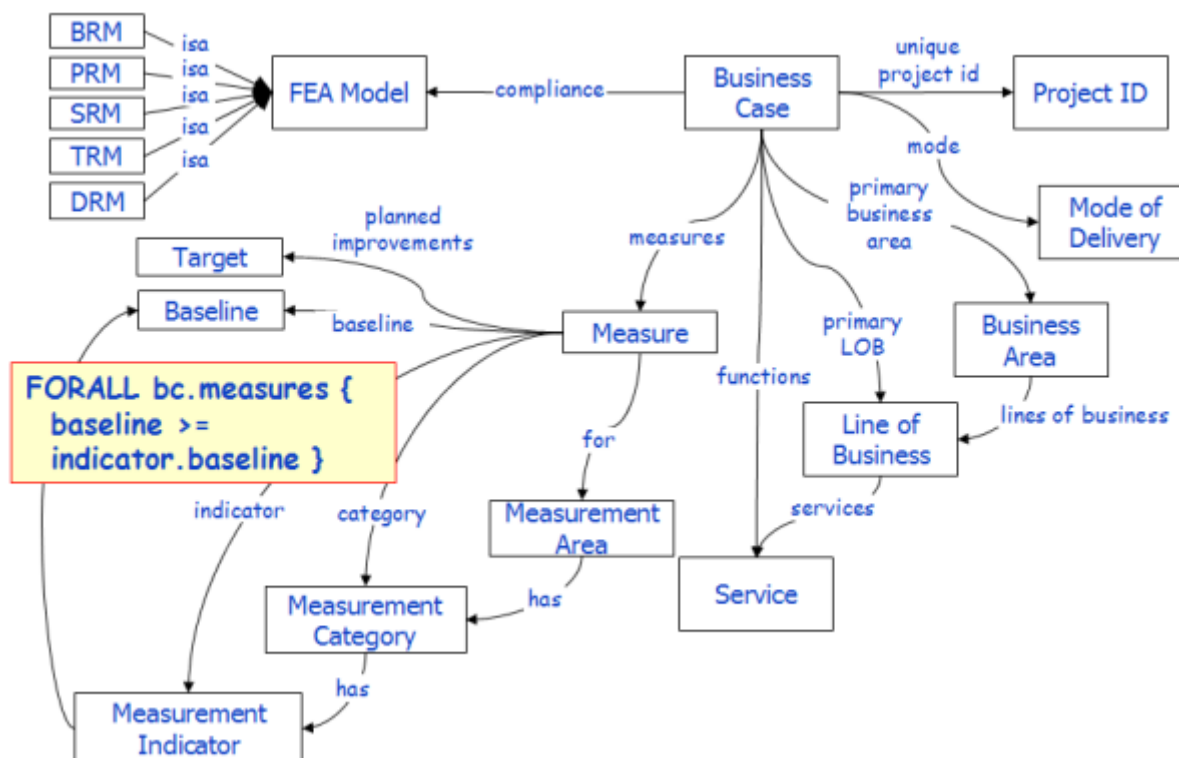
Another white paper from TopQuadrant, *Semantic Technology Briefing*, relates the [FEA](#) and E-Gov Initiatives to taxonomies and ontologies. Whereas a *taxonomy* represents hierarchical parent-child relationships with unnamed arcs, an *ontology* depicts a network of named connections from which new knowledge can be inferred by analyzing the connections between concepts. Figure 10 shows an ontology depicting a generic E-Gov Initiative.



SOURCE: [http://www.topquadrant.com/documents/TQ03\\_Semantic\\_Technology\\_Briefing.PDF](http://www.topquadrant.com/documents/TQ03_Semantic_Technology_Briefing.PDF) Used with permission of TopQuadrant.

**Figure 10. Part of the FEA Capabilities Manager Ontology Model**

Ontologies may also incorporate rules and constraints such as that shown in Figure 11 which checks baseline values of measurement indicators in a business case.



SOURCE: [http://www.topquadrant.com/documents/TQ03\\_Semantic\\_Technology\\_Briefing.PDF](http://www.topquadrant.com/documents/TQ03_Semantic_Technology_Briefing.PDF) Used with permission of TopQuadrant.

Figure 11. OMB 300 Ontology Model with Simple Rule

## 8. Summary

We have seen a number of legislative actions that have encouraged the adoption of XML in the US Federal Government. OMB and the CIOC have essentially required the use of XML in the emerging FEA. The impact of XML technology on the Presidential E-Gov Initiatives is clear. The vast majority of the 28 Executive Agencies are currently incorporating XML technology into their IT solution sets. While this paper cannot possibly cover every XML effort or even every agency, hopefully we have at least painted a broad picture illustrating the diversity of these federal government activities, ranging from E-Forms, to registries and repositories, to guidance documents, to myriad specialized applications. Readers familiar with other significant federal XML efforts are welcomed to contact the author.

## 9. Appendix: Early History of Markup Technologies in the US Federal Government<sup>11</sup>

Although some may think that XML markup is relatively new to the federal government, actually the concept of using markup to achieve intelligent documents is not new to the government at all. The Department of Defense, Internal Revenue Service, US Patent and Trade Office and the intelligence communities have been using SGML, the predecessor of XML, for over 15 years as the mechanism to achieve and maintain intelligence and longevity of documents.

DoD developed a set of standards under the umbrella title of the Continuous Acquisition and Lifecycle Support nee Computer Aided Logistics Support (CALS) initiative. SGML and Electronic Data Interchange (EDI) were two standards that were part of the CALS initiative. DoD established a working committee called the Electronic

<sup>11</sup>This historical section was written by Betty Harvey of Electronic Commerce Connection (ECC), <http://www.eccnet.com>. Used by permission.

Publishing Committee (EPC), made up of industry and military service representatives. In 1989, EPC developed the [CALSTable Model](#), which is the basis of the XML OASIS Table Model, used by every application that requires robust and complicated tables. The OASIS Table Model uses the same hierarchy, elements and attributes originally defined in CALS. Whereas the HTML table model is useful for scrollable displays such as desktops or laptops, it cannot accommodate the composition and pagination requirements of complex tables.

[IRS](#) was one of the first government agencies besides DoD to adopt SGML as the basis for their help documentation, as well as for their forms, provided in two formats, PDF and SGML (<http://www.irs.gov/form-spubs/lists/0,,id=97817,00.html>).

The US Patent and Trade Office ([USPTO](#)) began developing their electronic submission of patents process in SGML in the early 1990's. When XML became a standard in 1998, the agency migrated from SGML to XML. In 2000, USPTO developed a system to receive and allow patent submitters to submit their applications electronically using XML (<http://xml.coverpages.org/uspto.html>). USPTO in conjunction with their vendor, I4I, developed a customized XML patent applications authoring system using Microsoft Word that was offered free to patent attorneys and individuals applying for a patent.

## Acknowledgements

The author is grateful for the contributions of the following individuals:

*Owen Ambur*

For his support and leadership in the former XML Working Group of the US Federal CIO Council, as well as for the wealth of information he has collected at <http://xml.gov>.

*Betty Harvey*

For contributing the appendix entitled *Early History of Markup Technologies in the US Federal Government* which puts recent government work in the proper historical context.

*David Heiser*

For providing information about the significant work being done by IRS. David W. Heiser is Technical Advisor, IRS Tax Forms and Publications Division.

*Ralph Hodgson*

For providing the two FEA and OMB 300 ontology graphics used in the Semantic Web section.

*Brand Niemann*

For contributing one of his many presentations from which I was permitted to freely draw. Numerous Web services and Semantic Web presentations are available from his <http://web-services.gov> site.

*Rick Rogers*

For his leadership in the E-Forms for E-Gov pilot project under the Web Services Working Group.

*Marion Royal*

For his role in the GSA Office of Electronic Government (within the Office of Governmentwide Policy) which includes, among many other things, providing XML experts to various government agencies, such as GSA, EPA, and DON, and for his efforts in establishing XML registries.

## Bibliography

[Business Gateway] The Business Gateway will eventually be located at <http://Business.gov> although at the time of this writing, this URL was the home of the U.S. Business Advisor. Since the Business Gateway Initiative was in the process of active change, some of information presented in this paper may not be accurate as the initiative matures.

- [CCA] Clinger-Cohen Act of 1996. See [http://www.cio.gov/Documents/it\\_management\\_reform\\_act\\_Feb\\_1996.html](http://www.cio.gov/Documents/it_management_reform_act_Feb_1996.html) and [http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=1996\\_register&docid=fr19jy96-133.pdf](http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=1996_register&docid=fr19jy96-133.pdf).
- [CCTS] *UN/CEFACT - Core Components Technical Specification, Version 2.0* [August 2003] is available from [http://www.untmg.org/doc\\_ccwg.html](http://www.untmg.org/doc_ccwg.html).
- [CIOC XDG] The draft *XML Developers Guide* of the US Federal CIO Council is located at [http://xml.gov/documents/in\\_progress/developersguide.pdf](http://xml.gov/documents/in_progress/developersguide.pdf) [April 2002].
- [DoD XML and Metadata Registries] The DoD XML Registry is located at <http://diides.ncr.disa.mil/xmlreg/user/index.cfm>. The DoD Metadata Registry is located at <http://diides.ncr.disa.mil/mdreg/user/index.cfm>. The XML Schema that describes a registry submission is located at <http://diides.ncr.disa.mil/xmlreg/XSD/registry.xsd>. A sample submission package is available as a ZIP file from <http://diides.ncr.disa.mil/xmlreg/examples/BookSubmission.zip>. NOTE: A government-sponsored user account is necessary for access to most of the actual content of the registry.
- [DON XDG] The *XML Developers Guide, Version 1.1* of the Department of Navy is available from the XML Developers Library at <https://quickplace.hq.navy.mil/QuickPlace/navyxml/Main.nsf/>. Version 2.0 is expected before the end of 2003.
- [DRive] The Data Interchange Standards Association (DISA) Registry Initiative (DRive) is located at <http://www.disa.org/drive/>.
- [E-Forms] The home page for the E-Forms pilot is <http://www.fenestra.com/eforms/>. The related W3C XForms home is <http://www.w3.org/MarkUp/Forms/>. Other E-Forms efforts can be found at various agency sites, such as GSA, SBA, IRS, EPA, and NARA.
- [EGOV ACT] The E-Government Act of 2002. See <http://thomas.loc.gov/cgi-bin/bdquery/z?d107:hr2458> for the Act and <http://www.whitehouse.gov/news/releases/2002/12/20021217-5.html> for the press release and also [http://www.ssa.gov/legislation/legis\\_bulletin\\_112202a.html](http://www.ssa.gov/legislation/legis_bulletin_112202a.html).
- [EGOV INITS] President Bush's 25 E-Government Initiatives. See <http://www.whitehouse.gov/omb/egov/index.html> or <http://www.egov.gov>. An up-to-date chart of the status of the initiatives is located at [http://www.whitehouse.gov/omb/egov/downloads/E-Gov\\_Initiatives.pdf](http://www.whitehouse.gov/omb/egov/downloads/E-Gov_Initiatives.pdf). E-Gov Legislation, Reports and Resources are listed at [http://www.whitehouse.gov/omb/egov/about\\_leg.htm](http://www.whitehouse.gov/omb/egov/about_leg.htm).
- [EGOV XML] E-Government Act of 2002: Selected Portions of Interest to the XML Working Group; prepared by Owen Ambur, Fish and Wildlife Service. See <http://www.xml.gov/documents/completed/eGovXML.htm>.
- [EPA Registries] The EPA Environmental Data Registry is located at <http://www.epa.gov/edr/> and [http://oaspub.epa.gov/edr/xml\\$.startup](http://oaspub.epa.gov/edr/xml$.startup) with a search interface at [http://oaspub.epa.gov/edr/search\\$.startup](http://oaspub.epa.gov/edr/search$.startup). The EPA Environmental Information Exchange Network is located at <http://www.exchangenetwork.net/common/default.asp> with its XML Registry at [http://oaspub.epa.gov/emg/xmlsearch\\$.startup](http://oaspub.epa.gov/emg/xmlsearch$.startup).
- [EPA XDR] EPA's *XML Design Rules and Conventions for the Environmental Exchange Network* is located at [http://www.exchangenetwork.net/Documents/EP211S1\\_0\\_XML%20Design%20Rules%20Report.pdf](http://www.exchangenetwork.net/Documents/EP211S1_0_XML%20Design%20Rules%20Report.pdf) [draft, dated April 2003]. This long and comprehensive document presents many XML design issues, contains a helpful Summary of XML Rules, and covers several controversial issues such as namespace management. Version 1.0 was published in October 2003, but the URL was not available at the time of this writing.
- [ETravel Service] An overview of the eTravel Service (eTS), an E-Government initiative under the leadership of GSA, is located at <http://www.whitehouse.gov/omb/egov/internal/etravel.htm>. Considerably more detail about eTS and the other four GSA E-Government Initiatives can be found at <http://egov.gsa.gov>.
- [FEA] Federal Enterprise Architecture documentation including the five Reference Models is available from <http://www.feapmo.gov/default.asp> and [http://www.feapmo.gov/fea\\_downloads.asp](http://www.feapmo.gov/fea_downloads.asp). The XML Schema for the OMB 300 is available from [http://www.cio.gov/documents/fy2005\\_final\\_XML\\_schema.html](http://www.cio.gov/documents/fy2005_final_XML_schema.html).

- [GAO\_1] The GAO report entitled *National Preparedness: Technology and Information Sharing Challenges* [August 2002] is available at <http://www.gao.gov/new.items/d021048r.pdf>.
- [GAO\_2] The GAO report entitled *Electronic Government: Challenges to Effective Adoption of the Extensible Markup Language* [April 2002] is available at <http://www.gao.gov/new.items/d02327.pdf>.
- [GIDS] Generalized Instrument Design System, used by the Census Bureau. See the Fenestra case study located at <http://www.fenestra.com/default.asp?casestudies/gids.htm> and a paper from 2002 at [http://www.ideal-liance.org/papers/xml02/dx\\_xml02/papers/06-05-02/06-05-02.pdf](http://www.ideal-liance.org/papers/xml02/dx_xml02/papers/06-05-02/06-05-02.pdf).
- [GPEA] Government Paperwork Elimination Act. See <http://www.cdt.org/legislation/105th/digsig/govnopaper.html> and <http://www.whitehouse.gov/omb/fedreg/gpea2.html>.
- [GSA] Details about the five GSA E-Government Initiatives can be found at <http://egov.gsa.gov>.
- [House of Representatives XML] US House of Representatives XML and Legislative Documents are located at <http://xml.house.gov/>. DTD, XML Schema and XML instances are included. Works best with Internet Explorer.
- [IAE] An overview of Integrated Acquisitions Environment, an E-Government initiative under the leadership of GSA, is located at <http://www.whitehouse.gov/omb/egov/internal/acquisition.htm>. Considerably more detail about IAE and the other four GSA E-Government Initiatives can be found at <http://egov.gsa.gov>. Other relevant Web sites include: Business Partners Network at <http://www.BPN.gov>, Federal Business Opportunities at <http://www.FedBizOpps.gov>, Federal Technical Data Solution at <http://www.Fed-TeDS.gov>, Past Performance Information Retrieval System at <http://www.PPIRS.gov>, and Federal Procurement Data System "Next Generation" at <https://www.fpds.gov> (project/developer site: <http://www.fpds-ng.com>).
- [InfoPath] The InfoPath solution for OMB300 form is available from <http://www.microsoft.com/usa/government/fed/solutions/form300.asp> on Microsoft's US Government site, <http://www.microsoft.com/usa/government/fed/default.asp>. At the time of this writing, the home page for Microsoft's InfoPath 2003 preview is <http://www.microsoft.com/office/preview/infopath/default.asp>. The article *How XML Standards Are Used in Microsoft Office InfoPath 2003* was located at [http://msdn.microsoft.com/library/default.asp?url=/library/en-us/odc\\_ip2003\\_ta/html/odc\\_inXMLs.asp?frame=true](http://msdn.microsoft.com/library/default.asp?url=/library/en-us/odc_ip2003_ta/html/odc_inXMLs.asp?frame=true).
- [IRS] Searching the IRS Web site at <http://www.irs.gov> for XML results in quite a number of hits. The XML Schema are located at <http://www.irs.gov/pub/irs-schema/>; SGML/XML products are located at <http://www.irs.gov/pub/irs-sgml/>. See also Robin Cover's page, *IRS Establishes Online XML Developers' Forum for Employment Tax E-file System*, at <http://xml.coverpages.org/ni2002-09-30-b.html>. The *94x XML Developers' Forum - New Employment Tax e-file System* resources are located at <http://www.irs.gov/taxpros/providers/article/0,,id=97753,00.html>.
- [ISO 11179] International Organization for Standardization/International Electrotechnical Commission 11179 is a six-part data standardization and registry specification entitled *Information technology -- Specification and standardization of data elements*. Part 5: *Naming and identification principles for data elements* is frequently cited in US federal government XML guidelines. The ISO standard can be purchased from <http://www.iso.ch/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUM-BER=1758&ICS1=35&ICS2=40&ICS3=>.
- [JXDD] The home page of the Justice XML Data Model and Justice XML Data Dictionary (JXDD) is <http://it.ojp.gov/jxdm> and [http://it.ojp.gov/topic.jsp?topic\\_id=43](http://it.ojp.gov/topic.jsp?topic_id=43). The home page for Global Justice Information Sharing Initiative (Global) is <http://www.it.ojp.gov/global> or [http://it.ojp.gov/topic.jsp?topic\\_id=8](http://it.ojp.gov/topic.jsp?topic_id=8). See also the GTRI site at <http://justicexml.gtri.gatech.edu/index.html>.
- [Namespaces] GSA commissioned the preparation of *Recommended XML Namespace for Government Organizations* [draft, dated March 2003] located at [http://xml.gov/documents/in\\_progress/GS301L1\\_namespace.328.doc](http://xml.gov/documents/in_progress/GS301L1_namespace.328.doc). A new version may be available before the end of 2003.



[NARA] NARA's Records Management Metadata and Schema Project is located at [http://www.archives.gov/records\\_management/initiatives/schema\\_guidelines.html](http://www.archives.gov/records_management/initiatives/schema_guidelines.html). The NARA E-Records Management E-Government Initiative is described at [http://www.archives.gov/records\\_management/initiatives/erm\\_overview.html](http://www.archives.gov/records_management/initiatives/erm_overview.html).

[NIST XML Registry] The experimental GSA-NIST XML Registry Pilot is located at <http://xmlregistry.nist.gov:8080/index.jsp>.

[OMB 300] Office of Management and Budget Exhibit 300 (planning, budgeting, acquisition, and management of capital assets) for the current year is located at [http://www.whitehouse.gov/omb/circulars/a11/current\\_year/s300.pdf](http://www.whitehouse.gov/omb/circulars/a11/current_year/s300.pdf). OMB Exhibit 53 (information technology and E-Government) is located at [http://www.whitehouse.gov/omb/circulars/a11/current\\_year/s53.pdf](http://www.whitehouse.gov/omb/circulars/a11/current_year/s53.pdf). The parent document OMB Circular A-11 is located at <http://www.whitehouse.gov/omb/circulars/a11/03toc.html> for fiscal year 2004; this URL changes each year.

[Recreation One-Stop] This DOI E-Government Initiative is described at <http://www.recreation.gov>. Government Without Boundaries is located at <http://www.gwob.gov>. RecML, the Recreation Markup Language, is described at <http://www.recreation.gov/recmlintro.cfm> and <http://www.notes.fs.fed.us:81/recml>.

[Section 508] Section 508 of the Rehabilitation Act of 1998. See <http://www.section508.gov>. See also 36 CFR Part 1194, Electronic and Information Technology Accessibility Standards at <http://www.access-board.gov/sec508/508standards.htm>.

[UBL] Universal Business Language is an OASIS effort. See <http://www.oasis-open.org/committees/ubl>.

[VoiceXML] VoiceXML resources include <http://www.voicexml.org/>, <http://www.w3.org/Voice/> and <http://www.tell-me.com/>. The LEPC work is further described <http://xml.gov/documents/completed/VoiceXMLaward.htm>. See also the article *VoiceXML: A Publishing Standard for Accessibility* at <http://www.voicexmlreview.org/Apr2002/features/voicexml2.html>. The Disaster Management E-Gov Initiative Web site is <http://www.disasterhelp.gov>.

[Web Services and Registries Pilot] See *Web Services and Registries* at <http://web-services.gov/Web%20Services%20and%20Registries%20Pilot%20Proposal%20-%202007-22-03.ppt> and *A Government Enterprise Component Registry and Repository Using Native XML Database Technology* at <http://web-services.gov/Components%20Repository72203.ppt>. See also the Pilot Work Plan located at <http://web-services.gov/Web%20Services%20and%20Registries%20Pilot%20-%20PILOT%20PLAN.doc>.

[Web Services WG] The home page for the Web Services Working Group of the US Federal CIO Council is <http://web-services.gov/>. The working group was reformed as a Community of Practice in October 2003. The site contains a wealth of information about Web services and related technology such as the Semantic Web and E-Forms, including numerous presentations and links.

[XML Registries] Owen Ambur maintains a very comprehensive list of XML registries in use in US government agencies and several international registries as well. See <http://xml.gov/registries.asp>.

[XML WG] The home page for the XML Working Group of the US Federal CIO Council is <http://xml.gov/>. The working group was reformed as a Community of Practice in October 2003. The site contains a wealth of information about XML including numerous presentations and links.

## Glossary

ASC	Accredited Standards Committee
BRM	Business Reference Model
CALS	Continuous Acquisition and Lifecycle Support nee Computer Aided Logistics Support



---

CCA	Clinger-Cohen Act
CCR	Consolidated Component Repository
CEPPO	Chemical Emergency Preparedness and Prevention Office
CIOC	(the US Federal) Chief Information Officers Council
DHS	Department of Homeland Security
DISA	Defense Information Systems Agency
DOC	Department of Commerce
DoD	Department of Defense
DOE	Department of Energy
DOI	Department of the Interior
DOJ	Department of Justice
DON	Department of Navy
DOT	Department of Transportation
DRive	DISA Registry Initiative
DRM	Data and Information Reference Model
E-Gov	Electronic Government
EDI	Electronic Data Interchange
EDR	Environmental Data Registry
EPA	Environmental Protection Agency
ERM	electronic records management
eTS	E-Travel Service
FEA	Federal Enterprise Architecture
FEAPMO	Federal Enterprise Architecture Program Management Office
GAO	General Accounting Office
GIDS	Generalized Instrument Design System
GPEA	Government Paperwork Elimination Act
GSA	General Services Administration
IAE	Integrated Acquisition Environment
IRS	Internal Revenue Service
IT	Information Technology
JXDD	Justice XML Data Dictionary
JXRR	Justice XML Registry/Repository

LEPC	Local Emergency Planning Committee
NARA	National Archives and Records Administration
NASA	National Aeronautics and Space Administration
NIST	National Institute of Standards and Technology
NSA	National Security Agency
NSF	National Science Foundation
OMB	The Office of Management and Budget
OTA	Open Travel Alliance
PMO	Program Management Office
PRM	Performance Reference Model
RDF	Resource Description Framework
RecML	Recreation Markup Language
RFP	Request For Proposal
SAWG	Solutions Architects Working Group
SBA	Small Business Administration
SGML	Standard Generalized Markup Language
SRM	Service Component Reference Model
SSA	Social Security Administration
SVG	Scalable Vector Graphics
TRM	Technical Reference Model
UBL	Universal Business Language
UDDI	Universal Discovery, Description and Integration
UML	Unified Modeling Language
USDA	US Department of Agriculture
USPTO	US Patent and Trade Office
VA	Veterans Administration
VCS	Voluntary Consensus Standards
WSDL	Web Services Description Language
XSD	XML Schema

## Biography

**Kenneth Sall**  
Technical Architect/XML Specialist

SiloSmashers  
Contractor to GSA IAE Program Management Office  
Vienna  
United States of America  
KSall@SiloSmashers.com

Kenneth Sall of SiloSmashers, Inc. (<http://SiloSmashers.com>) served as XML consultant on the GSA Integrated Acquisition Environment E-Government Initiative. Ken has been closely tracking XML developments since July 1997 and has developed several XML training courses and numerous presentations for NASA and other government agencies. He created the extensive XML section of Web Developers Virtual Library (<http://WD-VL.Internet.com>). Addison-Wesley published Ken's comprehensive book, [XML Family of Specifications: A Practical Guide] in June 2002. His personal Web site (<http://kensall.com>) contains several useful XML resources including his unique [Big Picture of the XML Family of Specifications] (<http://kensall.com/big-picture/>), an imagemap gateway to all major XML technical specifications which also indicates their maturity and depicts interrelationships. Ken has been an active participant in the Federal CIO Council's XML Working Group (<http://xml.gov>), the Web Services Working Group (<http://web-services.gov>), and in three of the subteams of the E-Forms for E-Gov Pilots (<http://www.fenestra.com/eforms/>).