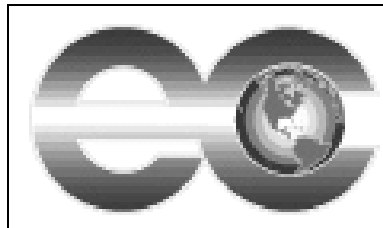


An Introduction to XML's Potential Use Within Government

December, 2000
NECCC



National Electronic Commerce Coordinating Council

The National Electronic Commerce Coordinating Council (NECCC) is an alliance of national state government associations dedicated to the advancement of electronic government within the states. The Council is comprised of **NASIRE** (a group representing the chief information officers of the states), the National Association of State Purchasing Officials (**NASPO**), the National Association of State Auditors, Comptrollers and Treasurers (**NASACT**) and the National Association of Secretaries of State (**NASS**). In addition to these voting members, other governmental and private organizations participate in an advisory capacity. These associations include the National Governors Association (**NGA**), the National Association of State Chief Administrators (**NASCA**), the Information Technology Association of America (**ITAA**) and the National Automated Clearing House (**NACHA**). The ITAA and the NACHA specifically represent private information technology companies and the financial services and technology industries.

EXECUTIVE BOARD

Chair: **Hon. J.D. Williams**, NASACT, Idaho State Controller

Vice Chair: **Richard B. Thompson**, NASPO, Maine Director of State Purchasing

Secretary/Treasurer: **Carolyn Purcell**, NASIRE, Executive Director, Texas Department of Information Resources

NECCC BOARD

NASIRE	George Boersma , CIO & Deputy Director, Michigan Department of Management & Budget Steve Kolodney , Director, Washington Department of Information Services
NASPO	Dave Ancell , Director, Office of Purchasing, Michigan Department of Management & Budget Joyce Murphy , Director, Division of Purchasing & Materials Management, Missouri Department of Administration,
NASS	Hon. J. Kenneth Blackwell , Ohio Secretary of State, Hon. Elaine Marshall , North Carolina Secretary of State Hon. Ron Thornburgh , Kansas Secretary of State
NASACT	Hon. David Heineman , Nebraska State Treasurer Hon. Thomas Wagner , Delaware Auditor of Accounts
NGA:	Thom Rubel , National Governors Association
NASCA	Pam Ahrens , Director, Idaho Department of Administration
ITAA	Basil Nikas , CEO, iNetPurchasing.Com
NACHA	William Kilmartin , Vice President, American Management Systems - AMS

NECCC CONTACT

Brandon W. Lenoir, Director • 202-624-5451 • e-mail: blenoir@nasact.org • web: www.ec3.org

Introduction

The move from the analogue state to the digital state has forced governments to rethink the way they do business. Integration of public sector services has created an environment that requires a cooperative, government-wide approach to service delivery instead of each organization working as an almost discrete entity.

Decision makers at every level of state and local government have wrestled with the question of how to transfer information between systems. Hyper-Text Mark-up Language (HTML) initially was thought to be the answer, but it provides only part of the solution.

The web today is based on HTML, and it is an effective tool for creating a “look” and “feel” for a portal. Unfortunately HTML falls short because it is unable to “tag” or label an object on one site so another site or system can recognize that same object. Enter eXtensible Mark-up Language, (XML).

XML provides a mechanism to label sets of data so they can be shared between systems. It is the underlying standard for systems communication and enables the transfer of data across agency lines. XML is already being used in many systems including cellular phones, palm pilots, and Windows applications.

The purpose of this document is to provide an overview of XML and the potential benefits of its use in the e-government transformation. It was prepared by the XML Project Group, a sub-group of the Integrated Services and Portals Workgroup of the NECCC. XML standards are examined, analysis of its potential benefits provided, and a discussion of ways government agencies are currently using XML to exchange data is given. In the conclusion the NECCC makes a call to all government decision makers to include XML in their planning efforts to create a seamless government.

WHAT IS XML?

XML is an open-standards-based technology, adopted by the World Wide Web Consortium (W3C), to enhance HTML for interchanging and processing data. Like HTML, XML is an abbreviated derivative of Standard Generalized Markup Language (SGML), the internationally agreed-upon standard meta-language.

There are four principle components that enable XML applications to process an XML document:

- The XML Document
- Document Type Definitions (DTD) or Schemas
- Processors and Parsers
- Style Sheets

THE XML DOCUMENT

The XML document is similar to an HTML document except that it contains custom tags. These tags are defined in a Document Type Definition (DTD). Following is a simple “business card” XML document:

```
<business_card>
```

```

<name>
  <first_name>Bob</first_name>
  <middle_int>B</middle_int>
  <last_name>Trout</last_name>
</name>
<title>STAC Guy</title>
<contact_methods>
  <phone>512-555-8160</phone>
  <fax>512-555-4759</fax>
  <e_mail>bob.trout@stac.state.tx.us</e_mail>
</contact_methods>
</business_card>

```

DOCUMENT TYPE DEFINITION

A **DTD** is where a set of custom tags, and attributes are defined. These are the rules that define and dictate how the custom tags relate to one another and how they can be used in the XML Document. The DTD for the preceding business card example may look something like this:

```

<!ELEMENT business_card(name, title, contact_methods)>

<!ELEMENT name(first_name, middle_int, last_name)>
<!ELEMENT first_name (#PCDATA)>
<!ELEMENT middle_int (#PCDATA)>
<!ELEMENT last_name (#PCDATA)>

<!ELEMENT title (#PCDATA)>

<!ELEMENT contact_methods(phone, fax, e_mail)>
<!ELEMENT phone (#PCDATA)>
<!ELEMENT fax (#PCDATA)>
<!ELEMENT e_mail (#PCDATA)>

```

DTDs can be either attached to the XML document or maintained in a specific location, such as a unique directory on a web server. If the location and syntax of these DTDs is agreed upon and shared by multiple partners, then this repository is known as a Namespace.

A **Namespace** is a method of identifying custom tags to avoid conflict. Namespaces provide the mechanics to identify that certain tags are associated with a particular entity and provide a standard way in which multiple entities can reduce naming conflicts.

PROCESSORS AND PARSERS

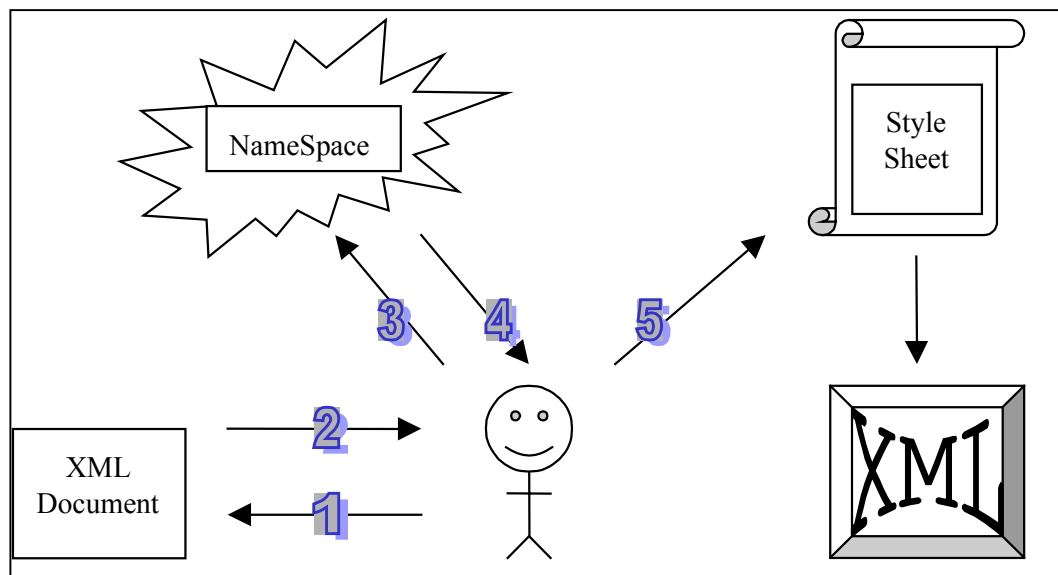
The XML processor, usually embedded in the application used to view the XML document, checks to make sure the XML file or document follows all the rules defined in the DTDs. This means the document has only one root element, all tags have matching start and closing tags, and there is no overlapping of tags.

The parser is responsible for reading the XML document and creating its document tree. This tree is a hierarchical structure of the content of the XML document and is once again defined in the DTD.

STYLE SHEETS

A style sheet is used to read the XML document and then associate the custom tags with particular display or presentation properties.

Following is a graphic representation of what an individual experiences using a web browser to review an XML document:



Step 1: The user requests an XML document

Step 2: The server hosting the XML document returns it to the user.

Step 3: Once the user has the XML document the DTDs are checked. The DTDs may be attached to the XML document or may be hosted in a defined Namespace.

Step 4: If the DTD resides in a Namespace it is retrieved. With the DTD the processor checks the document to make sure the XML is well formed and the parses are in the document. The browser does all of this without user intervention.

Step 5: If the document is well formed a Style Sheet is used to present the XML document's content to the user. The results are what is displayed in the user's browser.

THE STANDARDS

The leading standards body on XML is the World Wide Web Consortium (W3C), which recommended XML version 1.0 in February 1998. Since then several initiatives to

drive XML have emerged. RosettaNet, a global consortium dedicated to promoting open and common business processes, has defined a standard vocabulary to promote the inter-company exchange of data in the computer and electronics industry. CommerceNet eCo working group is developing a common XML framework to ensure interoperability across operating environments and across multiple industry standards. ASC, X12 XML/EDI group initiatives and UN/EDIFACT's simple EDI are working to adapt existing EDI standards to XML.

XML Standards:

Standard	Abbreviation	Status	Description
XML NameSpace	Namespace	Recommendation	Prevents overlap of names used by different software
SML Linking Language	Xlink	Working Draft	Flexible document link
Extensible Stylesheet Language	XSL	Working Draft	Flexible document processing
XSL Transformations	XSLT	Working Draft	Easy transformation of XML content
XML Schema	N/A	Working Draft	
FPML	FPML	Production	New XML based standard being developed to enable e-commerce in the financial derivatives sector
Biztalk	Microsoft led XML	Production	A Microsoft led effort to develop a set of guidelines for how to publish schemas in XML and how to use XML messages to integrate software programs together.
ebXML		Draft requirement	New general business standard being developed by the United Nations and Oasis consortium. The primary objective of ebXML is to lower the barrier of entry to electronic business in order to facilitate trade
XEDI		Production	This standard encapsulates traditional X12 and EDIFACT transactions in SML. Developed by the XEDI organization
EML/EDI		Under Development	Another Standard that encapsulates X12 and EDIFACT transactions in XML. XML/EDI Group is developing this standard.

CURRENT GOVERNMENT USES OF XML

Early collaborative projects already point to XML as “the next standard” for governmental information sharing. The enactment of the Health Insurance Portability and Accountability Act of 1996 (HIPAA) has led to vigorous activity within the health-care community to develop XML standards that will provide for improved portability and availability of group health coverage. Efforts to comply with HIPAA guidelines led to the formation of the Workgroup for Electronic Data Interchange (WEDI) <<http://www.wedi.org>>, a non-profit coalition including governmental, non-profit, and private-sector participants.

WEDI has organized an Emerging Technologies committee that has made XML one of its major initiatives. The committee issued a statement declaring, “The use of Web-based technologies to provide a universal access point for aggregating data from disparate systems seems a foregone conclusion. XML’s ability to facilitate the exchange of information between systems, coupled with the emerging applications to process XML

documents on the server, make it the obvious choice as the back-end framework for healthcare information systems.”¹

NASIRE, in its 1999 report, ***“Toward National Sharing of Governmental Information,”*** pointed to XML as a promising open standard for sharing dynamic structured information (i.e., dynamic documents) as a part of inter-agency or intergovernmental transactions. NASIRE states:

*The recent standardization of XML provides a robust extensible encapsulation method superior to any now in use; what is still lacking is widespread knowledge about and adoption of the new standard, both for ‘new-start’ sharing applications and re-work of existing applications. The development of document specifications for sharing requires substantial work at the national level...and then large amounts of painstaking work to complete the library of sharing documents.*²

That report was the result of a partnership between NASIRE and the U.S. Department of Justice (DOJ), Office of Justice Programs (OJP) and can be downloaded for free at <http://www.nasire.org/hotissues/justice>.

The justice community, with its history of nationwide – and even worldwide – information sharing, has already begun the painstaking task of developing shared XML documents. Legal XML (<http://www.legalxml.org>), a national non-profit organization founded in 1998, is comprised of volunteer members from private industry, non-profit organizations, government and academia. The mission of Legal XML is to develop open, non-proprietary technical standards for legal documents and related applications. The group has established numerous discipline-specific working groups (e.g., courts <http://www.courtxml.org>) and an Integrated Justice Working Group (IJWG) to facilitate information sharing across the justice system.

The IJWG has as its stated mission:

*...to identify and define the DTDs (document type definitions) for documents exchanged between courts, justice agencies, and key stakeholder agencies and organizations throughout the justice enterprise. The Integrated Justice Work Group will develop open, non-proprietary XML markup and related electronic interchange standards for documents routinely exchanged between federal, state, and local organizations in the justice environment.*³

Much of LegalXML’s document development process will likely complement work being done by SEARCH, the National Consortium for Justice Information and Statistics (<http://www.search.org>), which is compiling justice information “exchange points” as part of its ongoing integrated justice efforts. The exchange point study reports that:

¹¹ Bryan Caporlette, “XML offers a universal information platform for Healthcare Information Systems” n.d., <<http://www.wedi.org/htdocs/ETC/xml.htm>> (05 October 2000).

² NASIRE, “Toward National Sharing of Governmental Information” 1999, <<http://www.nasire.org/hotissues/justice/Fullrept.pdf>> (05 October 2000), 17.

³ LegalXML Integrated Justice Working Group “Charter” (v1.0) June 2000, 4.

*Agencies need to share critical information at key decision points throughout the justice process. Justice agencies conduct a series of what are essentially **conversations** at these key decision points in the justice process...*

At arrest, for example, the arresting agency typically sends certain information regarding the arrestee to the state criminal history records repository (for example, name, age, sex, race, driver's license number, electronic image of the arrestee's fingerprints, arrest charge information, etc.) to record the arrest transaction in the instant case...Every state, and to a lesser extent each jurisdiction within each state, is free, within certain constitutional and statutory boundaries, to construct and define their information exchange processes. Accordingly, there is considerable variation throughout the nation...Notwithstanding these and other variations, there appears to be significant consistency on key dimensions of information exchange – conversations – in justice processing throughout the nation.⁴

Every discipline of government could identify such conversations, both across agencies and between the local, state and federal levels. The federal government has begun investigating the potential uses of XML for enterprise-wide interoperability through the Federal CIO Council's XML Working Group. The council announced the formation of the group, which works under its Enterprise Interoperability and Emerging Information Technology Committee (EIEITC), saying, "eXtensible Markup Language (XML) embodies the potential to alleviate many of the interoperability problems associated with the sharing of data within and across organizations." The working group will undertake four activities: (1) Identify pertinent standards and best practices; (2) Establish partnerships with industry and public interest groups; (3) Establish partnerships with governmental communities of interest; and (4) Education and outreach.⁵

The time has come for governmental entities and their related IT associations to aggressively pursue enterprise-wide XML DTDs for information exchange. They should consider opportunities to plug into existing private-sector efforts such as the HR-XML Consortium (<http://www.hr-xml.org/>), a national, non-profit organization seeking "an e-commerce framework for human resources." This consortium offers a free membership to non-profits, which includes participation in technical working groups and voting rights.

Applying XML within government can yield greater benefits if government agencies inventory common horizontal and vertical data-exchange "conversations" and agree to standard DTDs. Public-private partnership can promote the development of XML-enabled software that supports the predominant document formats, saving governments from endless custom scripting to map variant DTDs to local standards.

⁴ SEARCH "Planning the Integration of Justice Information Systems: Developing Justice Information Exchange Points" March 2000, <<http://www.search.org/integration/pdf/ExchangePoints.pdf>> (05 October 2000), 5-6.

⁵ Federal CIO Council "Memorandum" 25 September 2000, <http://www.cio.gov/docs/XML_memo_to_CIOs_9-00.htm> (05 October 2000).

SMART PORTALS IN E-GOVERNMENT

"Smart Portals" will use XML to change the way governments manage and transfer their "business" information. Portals are a virtual meeting place for a community of businesses and entities and a complementary community of customers. HTML helped build first level portals which guide site visitors to information in an intuitive fashion. XML will create *Smart Portals* that aggregate and customize data across the web for presentation to the smart portal visitor.

Agents interact with the user and various "information marketplaces" to search for information described through meta-tags. Intelligent, mobile software agents can perform such "human" tasks as searching, querying, and negotiating, and can also extend their own capabilities by cooperating in various ways with other mediator agents to obtain additional search strategies. The objective in blending Agents and XML for Smart Portals is to develop new technologies that inter-operate with the old technologies for file manipulation, text processing, etc. An XML repository may contain one or more glossaries, each designed for use in a specific functional area. The repository would also include the DTDs, schema, stylesheets, code sets and pre-programmed components that applications can use to ensure consistency and maintainability of shared definitions and processes. The XML repository must be designed to accommodate accessibility requirements for both manual access (ADA requirements) and automated systems using standard application program interfaces. Repositories maintain data to enhance security and version control. For example: transactions executed in the health-care community must be able to identify what, which version, and when information was obtained from an XML repository.

The Federal CIO Council's XML Working Group is currently evaluating the requirements for a government XML repository, and has estimated development cost at between \$1-2 million, with annual maintenance cost of \$1million.

The potential benefits of using XML are:

- Provide for self-described transactions;
- Enhance workflow and document management functions;
- Interface with legacy systems;
- Reiterate the use of object-based documents;
- Concurrently provide object attributes and methods;
- Create a semantic web application rather than "batch" transactions;
- Ease of implementation;
- Acceptance by a number of vendors and partners;
- Provide e-government initiatives the path to move to a document-centric environment.

DOCUMENT-CENTRIC ENVIRONMENT

On the Internet electronic document-based management systems are integrating with relational database management systems (RDBMS). Middleware technology and

Universal Servers are combining fielded data with full text-based subsystems; database systems and document-oriented tools are supplementing each other and providing the end users with new paradigms for working with documents. Portal applications will use built-in web browser functions as their front-ends, so that developers will not need to write individual screens to view transactions such as an invoice or welfare record, but instead will use the integrated browser, which will *intelligently* "know" how to adapt for the transaction. State services, local governments and national e-commerce alliances need to collectively define DTDs for specific sectors and government services.

CONCLUSION

For decades IT professionals have recognized the distinction between the interoperability of hardware and software and the thornier problem related to the meaningful exchange of information. Technical difficulties fall away while those issues that require collaboration between different interest groups remain intractable. One of the ways that hardware and software vendors have accelerated their ability to interoperate has been to agree on an intermediate standard. Companies can develop products with unique qualities and characteristics for a particular niche, but the standard provides a pathway for their customers to transport the data to a software package used by a supplier or business partner.

Traditionally, databases have been custom built in silos, each reflecting the biases and peculiarities of its owner. Databases represent huge sunk costs for government. They also represent a barrier to sharing information because making changes can be very costly. Although XML is not a magic bullet, it affords the opportunity to overcome the barrier of rewriting proprietary databases. Electronic purchasing applications offer an excellent example of how XML is used. The suppliers, each of whom has a proprietary catalogue, agrees to an XML schema that allows an intermediary to aggregate the different catalogues and present them on the Web as a single purchasing resource for the industry. Some of the interest groups that have come together to develop a common intermediate standard for sharing information represent the following industries: banking, education, energy, healthcare and publishing.⁶

The imperative to provide a more integrated approach to governing and services is on the lips of nearly every elected official. XML should be factored into governments' plans for seamless government.

⁶ http://xml.org/xmlorg_registry/index.shtml

References:

1. **Digital Atlas 2001: Carto Vision 2001**, *Harsh Verma*, 16th INCA International Congress, INCA, Dec 16-19, 1996.
2. **Active Directory Services on Windows 2000**, *Microsoft*
3. **Introducing XML**, *Bruce Peat & David Webber*, *The XML/EDI Group*, August 1997
4. **Technology Perspectives for Information Overload**, *Erik Mettala*, *Defining the Advanced Technology Challenges of the Electronic Commerce Marketplace*, March 9-10, 1998, NIST, Gaithersburg, MD.
5. **Interoperable Infrastructures for Distributed Electronic Commerce**, *Shirley M. Hurwitz*, NIST
6. **XML and the Second-Generation Web**, *Jon Bosak and Tim Bray*, Scientific American

IS&P Workgroup Members

Al Sherwood	Utah	Keith Brown	Projectguides
Alan L. Weimer	Litton TASC	Laura Metzger	SAIC
Albert Powell	Ohio	Marc Cummings	ITStrategies
Ann Orlando	SAIC	Mark W.	Center for Digital
Art Phaup	Virginia	Struckman	Government
Bill Keip	Metamor	Marshall High	Idaho
BJ Bellamy	Kentucky	Matthew A.	Natcity
Bob Moriarty	SAIC	Stuczynski	
Brandon W. Lenoir	NECCC	Michael F. Teller	Lockheed
Bryan K. Thurston	PMP	Michele Grisham	CISCO
Carolyn Purcell	Texas	Mike Clair	AMS
Chris Dixon	AMR	Mike Nickolaus	Dyncorp
Curtis Clark	IBM	Mike Peterson	CACI
Dan Greenwood	Massachusetts	Milford Sprecher	SAP
Dave Moon	Utah	Milford Wright	Dunn & Bradstreet
Davey S. Thorton	Andersen Consulting	Naomi Marr	imco
David P. Gragan	Oracle	P. K. Agarwal	NIC
Dolores A. Seward	SAP	Pam Sutton	Nevada
Doug McCoy	EDS	Pat Keith	Texas
Doug Robinson	Kentucky	Pat Lundeen	Oregon
Dr. Harsh Verma	RSSoftware	Phyllis Richards	Idaho
Ed Leachman	Louisiana	Randy Cole	IBM
Gary Murad	Sapient	Randy Kitchen	Kentucky
Georgia K. Marsh	Illinois	Rick Wheeler	Andersen Consulting
Ginger Kadlec	Access Indiana	Robert Smith	Wolfenet
Graham	AMS	Robert Tyre	Andersen Consulting
McConnell		Sandy Wright	Dunn & Bradstreet
Jaqueline Rosales	Broadvision	Sheryl Chuhack	Illinois
John Kost	TRW	Stephen Lyon	IBM
Jon Allison	Egovnet	Steve Fricker	AMS
Judi Zito	Miami-Dade County FL	Sue Kidder	Wisconsin
Julie Holbrook	ProjectGuides	Tom Little	SCTcorp
Kara LaPierre	NIC		

