XML for RPG Programmers: An Introduction

OCEAN Technical Conference
Catch the Wave

Susan M. Gantner
susan.gantner @ partner400.com
www.partner400.com

Partner400
Your partner in AS/400 and iSeries Education

XML appears to be the future of data interchange. It provides a powerful way to describe and encode data that is system independent. Based on HTML, XML is quickly becoming the standard data interchange language for business-to-business applications.

In this introductory session you will learn what XML is, what the syntax looks like, how it can be used in your RPG applications, and how it will likely affect your applications in the future.

The author, Susan Gantner, is co-founder of Partner400, a firm specializing in customized education and mentoring services for AS/400 and iSeries developers. After a 15 year career with IBM, including several years at the Rochester and Toronto laboratories, Susan is now devoted to educating developers on techniques and technologies to extend and modernize their applications and development environments. This is done via on-site custom classes as well as conferences and user group events.

Together with her partner, Jon Paris, Susan authors regular technical articles for the IBM publication, eServer Magazine, iSeries edition, and the companion electronic newsletter, iSeries Extra. You may view articles in current and past issues and/or subscribe to the free newsletter or the magazine at: http://eservercomputing.com/iseries/.

Feel free to contact the author at: susan.gantner @ partner400.com and visit the Partner400 web site at www.partner400.com.

Notes

Agenda

What is XML?

- Relationship to HTML
- Writing XML: The Basics
  - Terminology, syntax and basic rules
  - Sister technologies
    - DTD, XML Schema, XSL

The Power of XML

- Examples of using it - now and future

iSeries and XML

- Tools and technologies specific to iSeries based on XML

What is XML?

eXtensible Markup Language

- A markup language for creating other markup languages!
  - A markup language uses tags inserted into information (such as HTML, UIM)
- XML allows the user to define a customized language to be used to format specific information
  - You create/define your own tags
- Defines the content and nature of the information -- not just how to display it in a browser
  - Some similarity to DDS
  - Standardization and related technologies give it more power
- Major focus is to provide a standard way to represent information
  - Standardization enhances information interchange

XML and W3C

XML standards defined by W3C

- Organized to establish specifications for web technologies ensuring the highest degree of utility and interoperability
- It creates and reviews specifications for XML and related technologies
- To see the latest standards information
  - http://www.w3.org/

A wide range of IT vendors provide extensive support for XML

- IBM, Microsoft, Sun, HP, Adobe, Oracle and many, many others

XML's Heritage

Both HTML and XML are derived from SGML

- Standard Generalized Markup Language
- Used in the document publishing industry - very complex
- HTML grew from the need to deliver information on the web
- XML addresses some of the limitations of HTML

Limitations of HTML?

- HTML tags are concerned with displaying information
  - Not with describing the content of the information
  - Fine for "computer to human" interactions
    - Less effective for computer to computer interactions

```
<p><b>Sally Smith</b><br>
Acme Computer Supplies
<br>
1234 Somewhere St.
<br>
Anytown, MN 55901
```

What is this?

XML and HTML

Partner400

HTML

- A tag based language
  - <tag> </tag> syntax
- Markup language for text
  - Focus on presentation
  - No meaning implied
- Fixed set of tags
- Loose syntax
  - Not case sensitive
  - End tags often not required
- Good for Computer to Human interaction
- Works with any browser

XML

- A tag based language
  - <tag> </tag> syntax
- Markup language for information
  - Focus on data structure
  - Data meaning implied
- Extensible-define your own tags
- Stringent syntax
  - Case sensitive
  - End tags always required
- Good for Computer to Computer interaction
- Works with few browsers "as is"
  - Can be transformed to HTML for presentation purposes

HTML Example

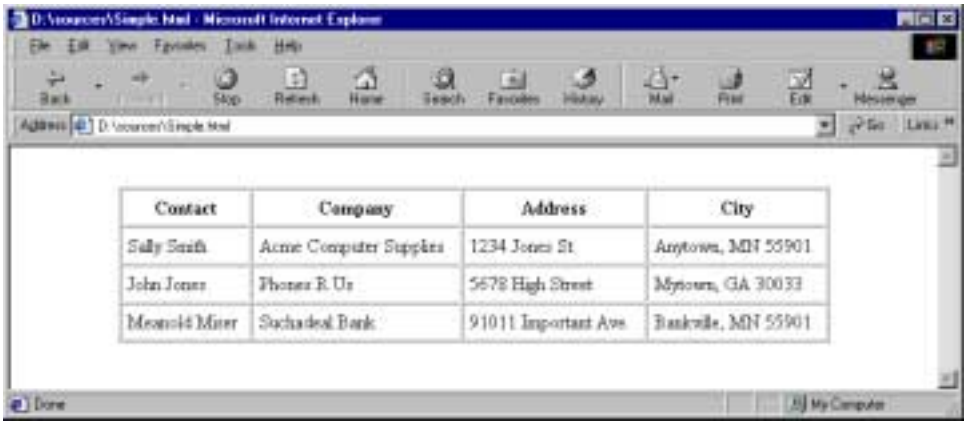
Partner400

```
<html>
<br><CENTER>
<table Border=1 CellSpacing=1 CellPadding=5 Width="600">
<th>Contact<th>Company<th>Address<th>City
<tr>
<td>Sally Smith</td>
<td>Acme Computer Supplies</td>
<td>1234 Jones St.</td>
<td>Anytown, MN 55901</td></tr>
<tr>
<td>John Jones</td>
<td>Phones R Us</td>
<td>5678 High Street</td>
<td>Mytown, GA 30033</td></tr>
<tr>
<td>Meanold Miser</td>
<td>Suchadeal Bank</td>
<td>91011 Important Ave.</td>
<td>Bankville, MN 55901</td></tr>
</table>
</html>
```

HTML Example, continued

HTML from previous page would look something like this in a browser

- Pretty, but the data has no meaning associated with it
- Only a presentation "view"



Contact	Company	Address	City
Sally Smith	Acme Computer Supplies	1234 Jones St	Anytown, MN 55901
John Jones	Phones R Us	5678 High Street	Mytown, GA 30033
Mearnold Miser	Suchadeal Bank	91011 Important Ave	Bankville, MN 55901

XML Example

```
<?xml version='1.0'?>
<Customers>
   <Customer>
     <Contact>Sally Smith</Contact>
     <Company>Acme Computer Supplies</Company>
     <Address>
       <Street>1234 Jones St.</Street>
       <City>Anytown</City>
       <State>MN</State>
       <Zip>55901</Zip>
     </Address>
   </Customer>
   <Customer>
     <Contact>John Jones</Contact>
     <Company>Phones R Us</Company>
     <Address>
       <Street>5678 High Street</Street>
       <City>Mytown</City>
       <State>GA</State>
       <Zip>30033</Zip>
     </Address>
   </Customer>
  <!-- One customer omitted from this example due to limited space -->
</Customers>
```

XML Example, continued

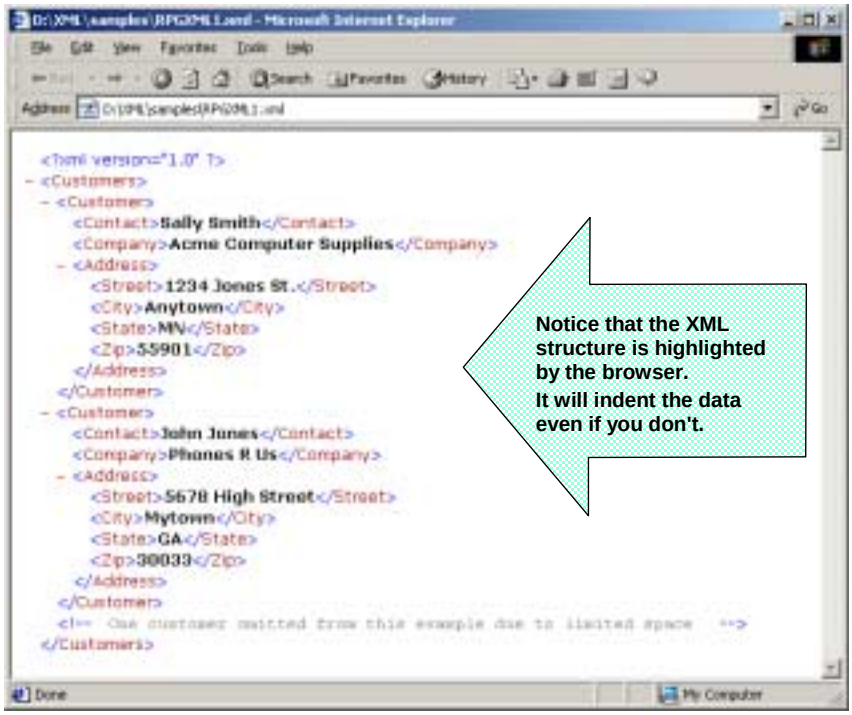
Note the descriptive names for the elements

- Emphasis is on the data content, not the presentation
  - Now we know what this data represents!
- Think how much easier it would be to search this data for "Jones" as a contact name
- How much easier is it to share this information with another application or system?
- XML document is roughly analogous to the data content in a physical file with field names embedded

But, what would it look like in a browser?

- Not much, as is
  - The browser will simply show the structure
- However, we can add XSL to give it "style"
  - XSL is Extensible Stylesheet Language
- We'll show you how this works later

What would it look like in a browser?



XML Terms

Elements

- The combination of a tag and its corresponding data
- Elements may contain other elements
  - The required Root element contains all other elements in the document
- XML documents consist of elements, such as:

<Street>1234 Jones St.</Street>

Attributes

- Additional information about an element

<Address type="shipping">

What are Schemas?

The definition of your customized markup language

- You define your own tags and the rules for using them
  - e.g., Element names, sequence, attributes: optional and required
- Schemas are used to **validate** an XML document
  - Does the document use the data elements correctly?
- Schemas are roughly analogous to the DDS for a PF
  - Except arrays and data structures can be defined as well as simple fields

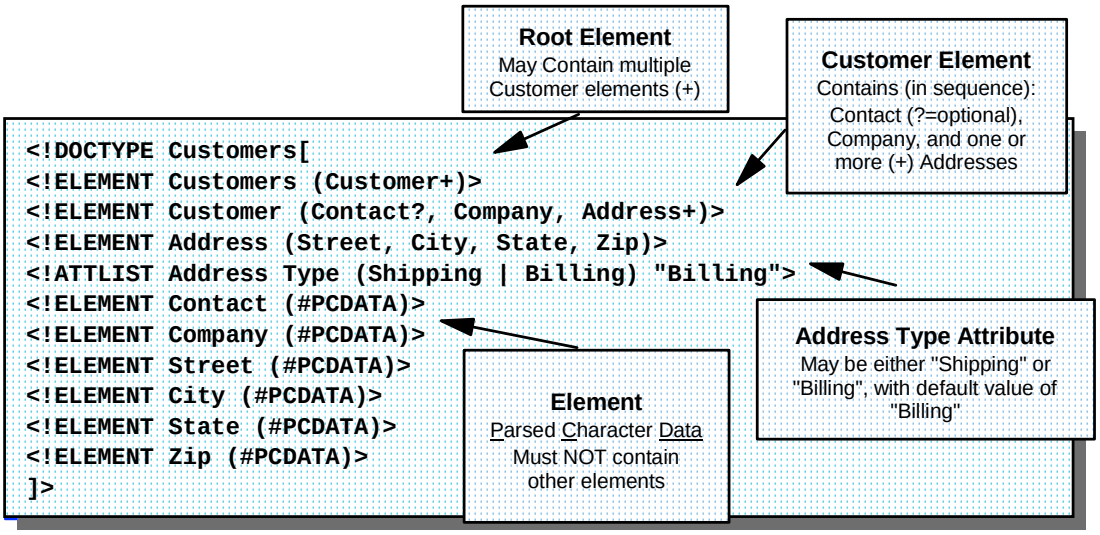
There are 2 ways to write schemas:

- DTD: Document Type Definition
  - Most widely used method to date
  - Ugly syntax, not consistent with XML syntax
  - Limitations in function
  - Will most likely be replaced over time with XML schemas
- XML Schema
  - Richer, more robust syntax consistent with XML documents
  - Developed by W3C
  - Standard recently finalized; therefore not widely used currently

DTD Example

A DTD for our previous XML document

- Defines the elements and attributes we can use in our document
- Note the beginning and ending square brackets [ ]
- May be coded inside the XML document or externally



What's Wrong with DTDs?

Awkward syntax inconsistent with XML syntax

Limited in function

- #PCDATA not a specific type of data
  - numeric, character, date, etc. do not exist as unique data types
- Cannot express some constraints easily, e.g.:
  - an element can occur 3 to 5 times
  - an element (zip code) is a 5 digit number with an optional hyphen and 4 additional digits

XML Schemas are designed to alleviate those problems

- Standard was only recently finalized by W3C
- For that reason, most current applications use DTDs

XML Schema Example

This code shows examples of XML Schema's advantages

- Ability to create new data types
- Constraints for minimum and maximum occurrences of elements

```
<!-- Note: This is a subset of the XML Schema for our sample XML -->
<xsd:element name='Address' type='StreetAddress'
<xsd:complexType name='StreetAddress' minOccurs='1' maxOccurs='2'>
 <xsd:sequence>
   <xsd:element name='Street' type='string'/>
   <xsd:element name='City' type='string'/>
   <xsd:element name='State' type='string'/>
   <xsd:element name='Zip' type='USZipCode'/>
 </xsd:sequence>
</xsd:complexType>
<xsd:simpleType name='USZipCode'>
 <xsd:restriction base='xsd:string'>
 <pattern value='[0-9]{5} (-[0-9]{4})?'/>
</xsd:simpleType>
```

Element
Built-in type 'string'

Element
User-created type 'USZipCode'

XML Document Rules

A "Well formed" XML document

- One that conforms to basic syntactical rules of XML
  - e.g., must contain at least 1 element, closing tags for all starting tags, no overlapping tags within nested elements
  - Similar to the type of checking done by an editor for your code
  - A document can be valid without a schema

A "Valid" XML document

- A well formed document that also conforms to the rules defined by a schema (such as a DTD)
  - e.g., only defined element names (tags) and attributes are used, nested elements adhere to rules, etc.
  - A valid document must have a schema
    - Either in the form of DTD or XML Schema
  - Similar to the type of checking done by a compiler
    - Except, you WROTE the rules for this compiler!

What is XSL?

eXtensible Stylesheet Language

- A language to transform XML documents

Typically, but not necessarily, for purposes of display

- Two parts:
  - XSLT (for Transformation)
    - Used to translate XML data into HTML for display in a browser
    - Often used to transform the same document for display to different types of output devices, such as cell phones, PDAs, etc. as well as browsers
    - Sometimes used simply to transform from one XML document to another XML document (e.g., similar data, different schemas)
  - XSL-FO (for Formatting Objects)
    - More focused on visual rendering of the XML document

XSL is roughly analogous to relatively simple program logic that re-formats data from an XML document either for display purposes or to write to a file in a different format

XSL Example

This XSL creates a simple HTML table from our XML data

```
<xsl:template match='/'>
<html><head><title>Customer List</title></head>
<body><CENTER>
<table Border='1' CellSpacing='1' CellPadding='5' Width='600'>
<th>Contact</th><th>Company</th><th>Address</th><th>City</th>
  <xsl:for-each select='Customers/Customer'>

     <tr><td><xsl:value-of select='Contact'/></td>
     <td><xsl:value-of select='Company'/></td>
     <td><xsl:value-of select='Address/Street'/></td>
     <td><xsl:value-of select='Address/City'/>,
     <xsl:value-of select='Address/State'/> -
     <xsl:value-of select='Address/Zip'/></td></tr>

  </xsl:for-each>
</table>
</CENTER>
</body></html>
</xsl:template>
```

Note that the XSL as shown on the previous page is not sufficient to produce the browser output. The full XSL is shown below (the additional lines are highlighted in **bold face**):

```
<?xml version='1.0'?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform" version="1.0">
<xsl:template match='/'>
<html><head><title>Customer List</title></head>
<body><CENTER>
<table Border='1' CellSpacing='1' CellPadding='5' Width='600'>
<th>Contact</th><th>Company</th><th>Address</th><th>City</th>
  <xsl:for-each select='Customers/Customer'>
     <tr><td><xsl:value-of select='Contact'/></td>
     <td><xsl:value-of select='Company'/></td>
     <td><xsl:value-of select='Address/Street'/></td>
     <td><xsl:value-of select='Address/City'/>,
     <xsl:value-of select='Address/State'/> -
     <xsl:value-of select='Address/Zip'/></td></tr>
  </xsl:for-each>
</table>
</CENTER>
</body></html>
</xsl:template>
</xsl:stylesheet>
```

In addition the following line is added to the XML in order to link it to the XSL file (the name of the file is highlighted in **bold**:)

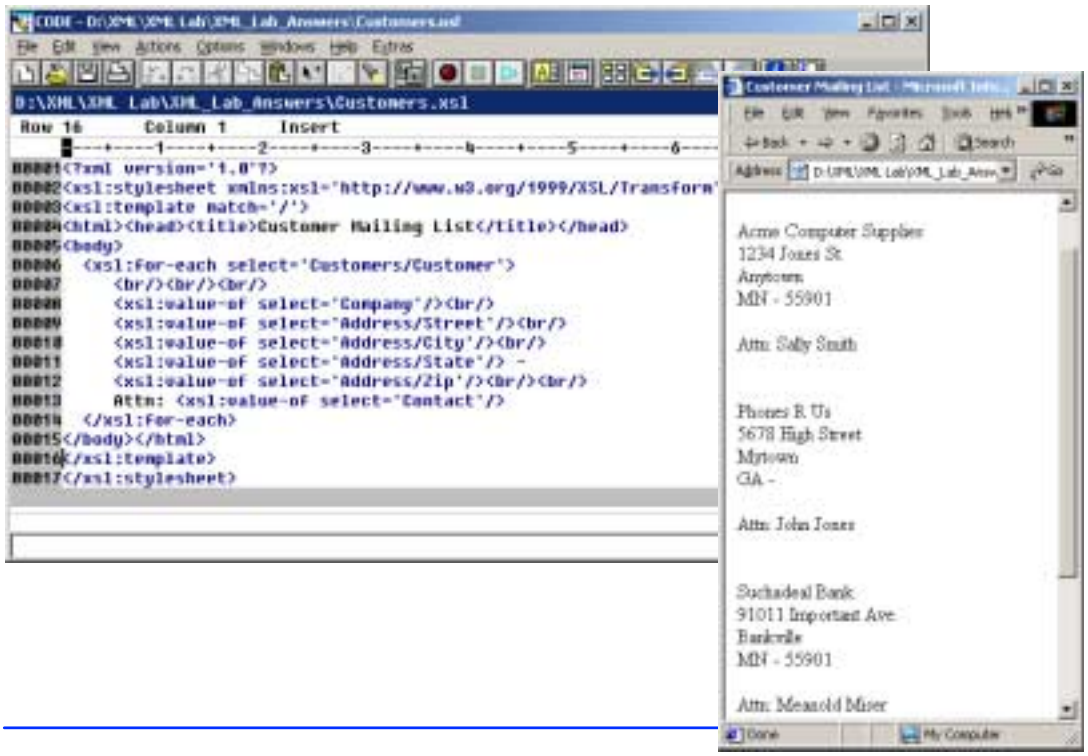
<?xml:stylesheet type="text/xsl" href="**RPGXML1.xsl**" ?>

XSL Example Results

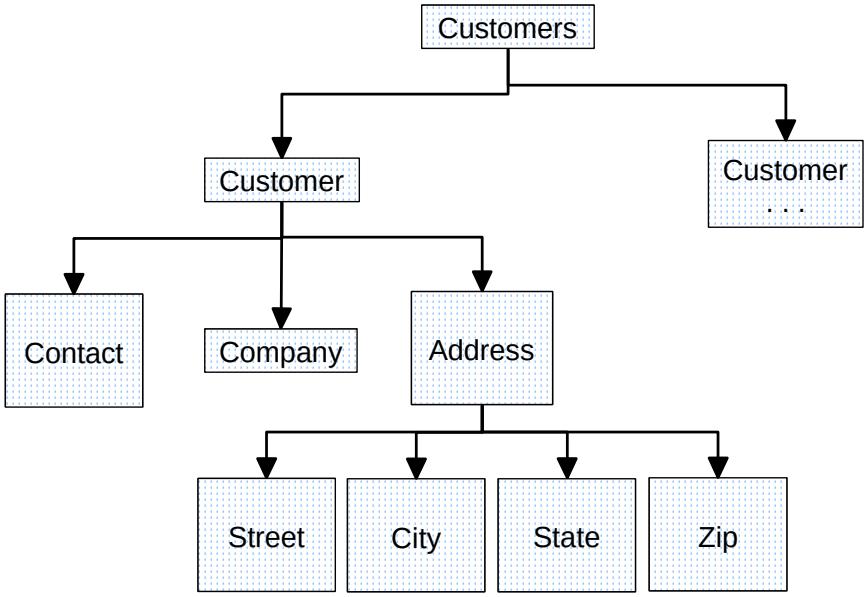
The XSL renders the XML data like this in a browser

- We could render the same XML data in many different ways with different XSLs

Different XSL - Different Result



XML Tree Structure



XML and RPG or COBOL Applications

If writing out new XML documents

- Easy enough to write XML syntax with the data using simple RPG/COBOL logic

If reading or updating existing XML documents

- XML's flexible syntax structure makes parsing/interpreting XML document structures very cumbersome with RPG/COBOL logic
- XML parsers are available - designed for use with Java or C++
  - IBM AlphaWorks has created APIs usable from RPG or COBOL
  - "XML Interface for RPG" Licensed Program 5733-XT1
- Two major types of parsers available
  - Document Object Model APIs
    - Creates tree structure in memory
    - Application has random access to all elements
    - Application can update or add to the XML document
  - Simple API for XML
    - Sequential access to elements
    - Input only - no update/add

How Can XML Be Used?

HTML replacement

- Multiple views of the same information
  - Single XML document, multiple stylesheets
    - For example by changing the stylesheets you can present the same data on a computer monitory, PDA, Cell-phone, etc. etc.
- Improved context based searching - "smart search"
  - Advanced search engine at IBM developerWorks/xml site is an example
  - Enables searching within a particular context
  - http://www.ibm.com/developer/xml/

How Can XML Be Used?

Standard Data Interchange Format

- Different organizations rarely standardize on a single set of tools
  - Work is required on both sides in order to exchange data
  - Even then, meaning sometimes gets lost in the translation
- New business to business data exchange formats based on XML
  - XML for product catalog, purchase order, invoice
  - Industries are declaring industry-standard XML protocols
  - Provides a bridge for sharing data between incompatible business systems
  - Could replace the need for some EDI-type applications
    - Less costly, more pervasive infrastructure
  - Insulates business applications

How Can XML Be Used?

Foundation for Pervasive Computing

- Standards for presenting content on mobile, pervasive devices
  - WML (Wireless Markup Language)
    - Part of a suite of application standards for cell phones, PDAs, other mobile devices
  - VoiceXML
    - Standard defined for presenting information presented by speech driven interfaces
  - Transcoding
    - Conversion of content on-the-fly to match device capabilities
    - XML and XSL are commonly used for transcoding application content

XML Support on iSeries

Partner400

iSeries has several XML-based technologies

- PDML - Panel Definition Markup Language
  - XML used to define GUI layout for Java applications
  - Runtime generation of PDML to Java/Swing GUI
  - Part of AS/400 Toolbox for Java
  - Was used to build Operations Navigator
- PCML - Program Call Markup Language
  - XML based language to define program interfaces
  - Part of AS/400 Toolbox for Java
  - Simplifies Java programs - especially when calling non-Java *PGMs
    - Handles complex relationships, such as varying length data, nested arrays
    - Provides data type translation and default values for input fields

XML Support on iSeries

Partner400

WebSphere Application Server

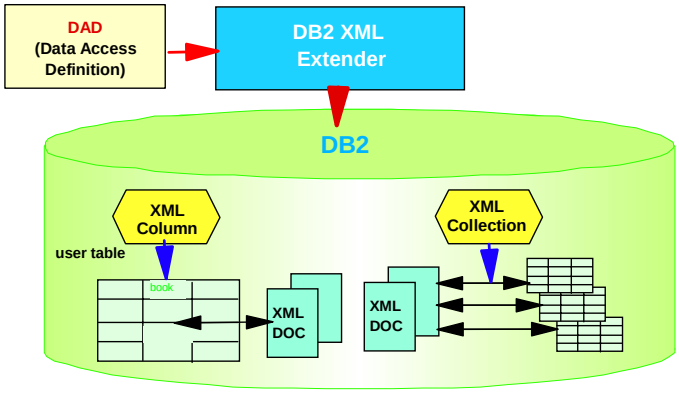
- Document Structure Services component contains features used in building application supporting XML data
  - xml4j - IBM Java-based XML parser
  - LotusXSL - XSL stylesheet runtime processor
  - Library of common DTDs
  - Sample servlets which support use of XML

XML Support on iSeries

XML Application Enablers

- XML Parsers
  - Allows applications to read, write and update XML data
  - Java (xml4j) and C++ (xml4c) parsers
  - Interfaces for using xml4c parser from RPG, COBOL and C
    - XML Toolkit for iSeries - Licensed program 5733-XT1
  - Parsers included in OS/400 at V5R1
    - Otherwise available for download at IBM alphaWorks site
    - http://www.alphaWorks.ibm.com/
- XSL Stylesheet runtime support
  - LotusXSL (Java) available from alphaWorks and Webshere App Server
  - Also provided with V5R1
- XML Extensions to DB2/UDB
  - Available (software product) at V5R1
  - Provides interchange between data in XML format and relational data
    - Two-way interchange supported

DB2 XML Extender



XML column

- Store and retrieve entire XML documents as DB2 column data
- XML data represented by XML column

XML collection

- Decompose XML document into a collection of relational tables
- Compose XML documents from a collection of relational tables

XML Futures

Partner400

Related standards continue to evolve

- XML Schema
  - Once standard is finalized, will eventually replace DTDs
- XML Query
  - XML query language syntax and semantics
- Industry DTD/Schema efforts
  - Collaborative efforts to converge on common DTDs for use in specific industries
- Web Services
  - Self contained, modular applications
    - Described and published using XML standards
    - Programmatically discovered and invoked over a network
  - Examples:
    - Stock quotes, credit card verification, integrated travel planning
  - Based on complimentary standards
    - SOAP - Invoke services across internet backbone
    - WSDL - Service interface definition
    - UDDI - Service registration and discovery

For more information . . .

Partner400

IBM developerWorks

- Latest news and information on XML
  - http://www.ibm.com/developer/xml

IBM alphaWorks

- Latest tools and enablers supporting XML
  - http://www.alphaWorks.ibm.com/

PartnerWorld for Developers, iSeries XML website

- Information and tutorials with iSeries flavor
  - http://www.iseries.ibm.com/developer/java/xml/index.html

W3C - XML standards and specifications

- Status and detail on various XML-related standards
  - http://www.w3.org/XML

XML.org - information on XML standards, tools

- Repository for industry standard DTDs shared across companies
  - http://www.xml.org/

Useful Reading, Tutorials, etc.

XML for the World Wide Web - Elizabeth Castro

- Low priced, easy to read general purpose introduction.
  - Well supported by website with downloadable examples
- Peachpit Press - www.peachpit.com
  - ISBN 0 - 201 - 71098 - 6

XSLT - Doug Tidwell

- Not a light read but a comprehensive and entertaining treatment by one of IBM's top XML "apostles"
- O'Reilly - www.oreilly.com
  - ISBN 0 - 596 - 00053 - 7

W3Schools.com - "the best things in life ARE free"

- Tutorials on all aspects of XML, XSL, and more
  - www.w3schools.com

ZVON.org - "The Guide to the XML Galaxy"

- Tutorials, links and more
  - www.zvon.org