XML (eXtensible Markup Language)

- Aniket Prabhune:

- Josh Steele:
  - Schemas, XSL/XSLT, Linking Languages (Xlink, XPath, Xpointer), XML Uses (MathML, XML-Data, XHTML, WIDL), Example
XML

- "The eXtensible Markup Language (XML) is the universal format for structured documents and data on the Web." – http://www.w3.org/XML/


- XML is a cross-platform, software and hardware independent, tool for transmitting information.

[PL00]
Origin of XML

- Development of XML started in 1996
- W3C Recommendation in 1998
- New but not immature!
- SGML an ISO standard since 1986
- HTML development started in 1990
- XML: Best parts of SGML guided by HTML
Need for XML

- XML is also a “document description meta-language” like SGML
- SGML is “more powerful ancestor” of XML
- SGML: difficult to learn and use, expensive
- XML: easy to learn and use, less expensive

[PL00]
XML Features

- XML is for structuring data
  - Spreadsheets, address books, financial transactions, technical drawings, etc
  - Set of rules
  - Not a programming language
  - Makes it easy for computer to generate and read data and ensure that the data structure is unambiguous
XML Features

- XML looks a bit like HTML but is different!
  - Both are markup languages that use tags and attributes
  - XML – not a replacement for HTML
    - XML: describe data – what data is?
    - HTML: display data – how data looks?
  - Ability to define own tags, attributes and document structure
  - Rules for XML are very strict
XML Features

- **XML is a text format**
  - Data Storage: Binary or Text
  - Text: + Easy to refer
    + Easy to debug

- **XML is Verbose by Design**
  - XML files are larger than binary formats
  - Inexpensive Disk space, compression programs, communication protocols
XML Features

- XML is a family of technologies.
  - XML 1.0
  - Xlink
  - XPointer and XFragments
  - CSS (Cascaded Style Sheet)
  - XSL/XSLT
  - DOM
  - XML Schemas 1 and 2
XML Features

- **XML is modular**
  - XML allows you to define a new document format by combining and reusing other formats

- **XML Namespaces:**
  - to eliminate name confusions
  - Collection of element type and attribute names
  - Identified by a unique name: URI
XML Features

- XML is well-supported, license-free and platform-independent
  - Large and growing community of tools and engineers experienced in the technology
  - License-free
  - Vendor independent
Information Exchange Models

Traditional:

Client (Browser)
Renders HTML for display purpose

HTML

Web Server
Sends information formatted in HTML so the client can display it

Business logic database

XML:

Client (Browser)
renders XSL/XML for display and can perform further tasks (e.g. sorting, calculations, etc)

XML/XSL

Web Server
Sends information in raw XML and (one time) includes XSL style sheet for formatting purposes to the client

Business logic database

[CH99]
Document Models

Traditional:

Information
Structure
Display

XML:

Information
Structure
Display

[CH99]
A Typical XML System

XML Document (Content)

XML DTD (Rules)

(optional)

XML Parser (Conformity)

XML Application

[CH99]
XML Document (Content)

- Actual Data to be processed
- Rich description of information using XML syntax
- Based on entities containing:
  - Content: Actual Information (author of book, price of book, number of pages, etc)
  - Content is encased in markup
DTD (Document Type Definition)

- Ensuring the structure of data
- Piece of code that defines the allowable structures in an XML Document
- Advantages of using DTD:
  - Check your XML document for validity
  - Share your data easily over the web
- Valid XML documents
- Well-formed XML documents

[CH99]
XML Parser

- A software engine that performs the actual check on the data to make sure:
  - It is syntactically correct (well formed)
  - It conforms to the DTD (valid) if you choose to include a DTD
- Replicates the structure of the information in memory, ready for customized processing through an XML application
- Included with Most Browsers
How a Parser Interprets XML?

XML Document

DTD (optional)

Well formed?

yes → Further Processing

no → Issue Warning/Stop Processing

DTD ?

yes → Valid ?

no → Issue Warning/Stop Processing

Valid ?

yes → Further Processing

no → Issue Warning/Stop Processing

[CH99]
XML Applications

- What you make as a programmer!
- Typically processes information encased in XML Documents
  - E-Commerce
  - Online Banking
  - Web Services
  - Creating other markup languages
  - Advanced Search Engines
  - Agents
  - Almost Anything!
Example
Example (DTD)

```xml
<<my_structure.dtd>>

<!ELEMENT INVENTORY (BOOK) +>
<!ELEMENT BOOK (TITLE, AUTHOR, PAGES)>
<!ATTLIST BOOK InStock (yes | no) #REQUIRED>
<!ELEMENT TITLE (#PCDATA)>
<!ELEMENT AUTHOR (#PCDATA)>
<!ELEMENT PAGES (#PCDATA)>
```
Example (XML Document)

<?xml version="1.0"?>
<!DOCTYPE INVENTORY SYSTEM "my_structure.dtd">
<?xml-stylesheet type="text/css" href = "my_display.css"?>
<!-- Beginning of Document Body -->
<INVENTORY>
  <BOOK InStock="yes">
    <TITLE>TCP/IP</TITLE>
    <AUTHOR>Comer</AUTHOR>
    <PAGES>245</PAGES>
  </BOOK>
  <BOOK InStock="no">
    ...
  </BOOK>
</INVENTORY>
<!-- End of Document Body -->
Example (CSS)

```
<<my_display.css>>

TITLE
{
  display:block;
  margin-top:12pt;
  font-size:20pt;
  font-style:italic;
  color: Blue;
}

AUTHOR
{
  display: block;
  margin-left:20pt;
  color: Red;
  font-size:20pt
  font-weight: bold;

  . . .
}
```
TCP/IP
Comer
245

Moby Dick
Herman Melville
223
References

- Information about XML: [http://www.w3.org/XML/](http://www.w3.org/XML/)
- XML 1.0 Recommendation: [www.w3.org/TR/REC-xml](http://www.w3.org/TR/REC-xml)
- Specific articles on XML: [www.xml.com](http://www.xml.com)
- [TB00] Tittel, E., Boumphrey, F., XML for Dummies, IDG, 2000

XML Tutorials:
- [http://wdvl.com/Authoring/Languages/XML/](http://wdvl.com/Authoring/Languages/XML/)
- [http://www.w3schools.com/xml/](http://www.w3schools.com/xml/)

XML Mailing Lists:
- [www.w3.org/XML/#discussion](http://www.w3.org/XML/#discussion)
- [www.oasis-open.org/cover/lists.html#discussionLists](http://www.oasis-open.org/cover/lists.html#discussionLists)
XML JOKE !!!!!!!

- When should you use XML?
- When you want a *buzzword* on your resume!