XML, XSL, XSLT, XPath, XPointer, BEEP and more
<XML> in detail
Definition

• **XML**: eXtensible Markup Language
  − A language for defining structured documents
  − Broad concept of what a “document” is (more general)

• **Who defines XML?**
  − A standard defined by the W3C (World Wide Web Consortium)
  − Brings together businesses and researchers

• **Documents are edited in text form combining data and tags**
  − Tags are written just like in HTML

<mytag>data</mytag>
XML = SGML + Web

- SGML
  - Created at the end of the 1970s based on work by Charles Goldfarb
  - Too complex for the web
- XML respects the SGML philosophy (distinction between data and presentation)
- XML is more flexible than SGML
  - The syntax is simpler (a subset of SGML)
  - Documents can be validated, or checked for being “well constructed”
- XML is richer than SGML
  - DSSSL and HyTime (SGML links) have not had the expected following
  - Hypertext links (XLL = XLink + XPointer)
  - Style sheets (XSL)
- There is an XML-compatible version of HTML: XHTML
Structured documents: definition

- A structured document consists of distinct logical elements.
- The composition of these elements defines a characteristic tree structure for the document.
- Bolstered by the distinction between:
  - The document's logical structure: its division into elements
  - The document's physical structure: saved in several files.
- The SGML standard was designed to allow for standardised description of logical structures.
Example of logical structure

```xml
<INVOICE>
  <CLIENT>
    <NAME>Poelman</NAME>
    <ADDRESS>43 Mercer NY, NY 1009</ADDRESS>
  </CLIENT>
  <LINE>
    <PRODUCT>single room</PRODUCT>
    <PRICE>340</PRICE>
    <QUANTITY>2 nights</QUANTITY>
  </LINE>
</INVOICE>
```

Hotel de la NYU

Greenwich Village, NY

Invoice n° 99005012

Addressed to:
Poelman
43 Mercer
NY, NY 1009

Payment by cheque upon receipt

<table>
<thead>
<tr>
<th>description</th>
<th>PU HT</th>
<th>Quantity</th>
<th>Subtotal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single room</td>
<td>340</td>
<td>2 nights</td>
<td>680</td>
</tr>
<tr>
<td>total</td>
<td></td>
<td></td>
<td>680</td>
</tr>
</tbody>
</table>
What is a DTD?

• Using a DTD makes it possible to check if a document is valid (that is, whether it follows its DTD)
  – Guarantees an exact structure before processing

• Explicitly describing a document’s logical structure in the form of a DTD makes it possible to:
  – Offer multiple presentations of a document
  – Analyse a document’s structure in order to “understand” it
  – Process the information it contains

• DTDs are a legacy of SGML
What is XML?

What is a DTD?

Example:

```
<?xml version='1.0'?>
<!DOCTYPE INVOICE [
<!ELEMENT INVOICE(CLIENT,(LINE)+)>
<!ELEMENT CLIENT (NAME,ADDRESS) >
<!ELEMENT NAME (#PCDATA) >
<!ELEMENT ADDRESS (#PCDATA) >
<!ELEMENT LINE (PRODUCT,QUANTITY,PRICE) >
<!ELEMENT PRODUCT (#PCDATA) >
<!ELEMENT QUANTITY (#PCDATA) >
<!ELEMENT PRICE (#PCDATA) >
]>
```

Invoice.xml
Inconveniences of a DTD

- Untyped elements (everything is #PCDATA)
- Complex syntax without tags (XML)
- No concept of inheritance for enriching or redefining elements
  - Concepts dealing with external physical entities only (inclusion of text)
- DTD history linked to SGML
- No concept of predefined types

- The solution: Schemas!
Why a schema?

1.1 Purpose

“The purpose of XML Schema: Structures is to define the nature of XML schemas and their component parts, provide an inventory of XML markup constructs with which to represent schemas, and define the application of schemas to XML documents.

The purpose of an XML Schema: Structures schema is to define and describe a class of XML documents by using schema components to constrain and document the meaning, usage and relationships of their constituent parts: datatypes, elements and their content and attributes and their values. Schemas may also provide for the specification of additional document information, such as normalization and defaulting of attribute and element values. Schemas have facilities for self-documentation. Thus, XML Schema: Structures can be used to define, describe and catalogue XML vocabularies for classes of XML documents.”

W3C Recommendation

- DTDs were too limited.
- No datatypes, everything was a string.
- DTD weren’t written in XML format and difficult to read and understand.
What is a schema?

• Example:

```xml
<xsd:schema>
  <xsd:element name="Invoice">
    <xsd:complexType>
      <xsd:sequence>
        <xsd:element name="Client">
          <xsd:complexType>
            <xsd:sequence>
              <xsd:element name="Name"/>
              <xsd:element name="Address"/>
            </xsd:sequence>
          </xsd:complexType>
        </xsd:element>
        <xsd:element name="Line" maxOccurs="15">
          <xsd:complexType>
            <xsd:sequence maxOccurs="unbounded">
              <xsd:element name="Product"/>
              <xsd:element name="Quantity"/>
              <xsd:element name="Price"/>
            </xsd:sequence>
          </xsd:complexType>
        </xsd:element>
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
</xsd:schema>
```

Invoice.xsd
The solutions schemas offer

- The XML Schema predefines various types (integer, date, URI, ...) It is possible to create custom data types
- XML schema takes complete control of namespaces I.e. name collisions are avoided e.g. two schemas may both have an element called Address without forcing a redefinition or ambiguity. Instead you can specify the namespace in front of the element. So you could have:
  
  `customerSchema:Address and supplierSchema:Address`.
- The ideas behind XML schema are very similar to objects (inheritance, abstract types, extension, aggregation)
- XML schema is an XML application (similar syntax)
- XML Schemas are usable as a platform and language independent IDL (Interface Definition Language) with extensive capabilities (Beyond CORBA IDL and Java IDL).
Well-formed documents and valid documents

• A well-formed XML document is a document that follows the rules for good construction listed in the specification, including:
  • Having all child elements completely enclosed in parent elements
    – One single root element in the document containing all others (an XML document is a tree)
    – A closing tag for each opening tag
• A valid XML document is a well-formed document that also respects the constraints defined in a DTD or a schema
Designing Schemas

• Schemas define the information systems Data Model
  – Abstracting the entities to be represented: contract, order form, brochure, product invoice, manual, catalogue, CV
  – Defining relationships between these entities

• Schemas quality is fundamental for
  – Workflow efficiency and flexibility
  – Expandability and integration of other data sources

• Tools are needed for
  – Designing the data model:
    • UML modeling
    • Database schemas
  – generating Schemas from analysis models
  – managing Schemas (sharing Schemas, version control)
XML schema

- A much better way of defining what an XML document must be structured as that the DTD was.
- It's an XML document itself.
- A template that defines the structure, order, and content of any XML documents that reference it.
- XML tag/value pairs can have a datatype associated with them. Thus enforcing validation of the syntax, structure and semantic content of an XML doc.
- Predefined datatypes include:
  - Byte
  - Short
  - Long
  - Float
  - Date
  - Time
  - String
  - Period,
  - Double float
  - Boolean
  - Language
  - Decimal
  - Others ...

- You can extend predefined datatypes and create your own!
Advanced XML & Schema terminology

From the W3C primer:

- **Definition:** Schema component is the generic term for the building blocks that comprise the abstract data model of the schema. An XML Schema is a set of schema components. There are 13 kinds of component in all, falling into three groups. The primary components, which may (type definitions) or must (element and attribute declarations) have names are as follows:
  - Simple type definitions
  - Complex type definitions
  - Attribute declarations
  - Element declarations

- The secondary components, which must have names, are as follows:
  - Attribute group definitions
  - Identity-constraint definitions
  - Model group definitions
  - Notation declarations

- Finally, the "helper" components provide small parts of other components; they are not independent of their context:
  - Annotations
  - Model groups
  - Particles
  - Wildcards
  - Attribute Uses
Advanced XML & Schema terminology

- **Schema** – Defines the structure that an XML doc must match. Like a class definition in OOP or structure of a table in DBMS terminology. Files have a .xsd extension.

- **Document** – the instance of a schema. Like an object is to a class in OOP. Or the actual table data vs. the definition in the DB schema. Files have a .xml extension. Documents do not have to reference a schema (or DTD) to be well formed but must reference one or more schemas to be valid.

- **Elements** - `<Tag value</Tag>` pair, contains one or more sub-elements and/or attributes. Like a class in OOP. Can contain 0,1, .. Inf sub-elements. Sub elements can contain sub-sub-elements, and so on. A node in the tree that is represented by XML. Elements have a tag, a closing tag, 0,1, .. n attributes, 0,1, .. n value(s), 0,1, .. n sub elements.

- **Attributes** - `<Tag MyAttribute="blue" Value</Tag>` An aspect of an element that only has one instance associated with that element. Ex: `<Person hair_color="brown" Mozart</person>`
Advanced XML & Schema terminology

• **simpleType** – an atomic datatype. ex: string, short, long, date, binary, …
  Ex:
  
  ```xml
  <xsd:element name="USPrice" type="xsd:decimal"/>
  ```

• **complexType** – an aggregate datatype. Created in your schema, a sequence of one or more Elements (complexTypes, simpleTypes, elements and/or attributes. Reusable definition of a datastructure. Like a `Struct` in C. 
  ex:
  
  ```xml
  <xsd:complexType name="USAddress">
    <xsd:sequence>
      <xsd:element name="name" type="xsd:string"/>
      <xsd:element name="street" type="xsd:string"/>
      <xsd:element name="city" type="xsd:string"/>
      <xsd:element name="state" type="xsd:string"/>
      <xsd:element name="zip" type="xsd:decimal"/>
    </xsd:sequence>
    <xsd:attribute name="country" type="xsd:NMTOKEN" fixed="US"/>
  </xsd:complexType>
  ```

• **Sequences** – elements that must appear in that specific order in the doc.
• **Choice** – only one of the sub elements is allowed in a doc.
• **Constraints** – minOccurs (default = 1), maxOccurs, (default = 1) optional, required, prohibited, fixed, default. Apply to both elements and attributes.
• **Facets** – aspects of an element or attribute. Ex: white space treatment, precision, …
• **Patterns** – limits the allowable values by a regular expression syntax e.g. `&#XXX`,
• **Enumerations** – a limited list of available choices for an element or attribute. Ex: US State abbreviations AL, TX, FL, …
XML Spy user interface
PersonDummy.xml

<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XML Spy v3.5 NT (http://www.xmlspy.com) by Logan Poelman (Valtech) -->
<person>
  <name>ABCD</name>
  <address>ABCD</address>
  <phone>999-999-999-9999</phone>
  <email>ABCD</email>
  <DOB>1999-12-12</DOB>
  <sex>Male</sex>
  <sex>Female</sex>
  <sex>Other</sex>
</person>
Schema generated from PersonDummy.xml by XML spy
Two different ways to structure an XML file that contains information about a person

```xml
<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XML Spy v3.5 NT (http://www.xmlspy.com) by Logan Poelmann -->
<person>
  <name>ABCD</name>
  <address>ABCD</address>
  <phone>999-999-999-9999</phone>
  <email>ABCD</email>
  <DOB>1999-12-12</DOB>
  <sex>Male</sex>
  <sex>Female</sex>
  <sex>Other</sex>
</person>
```

```xml
<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XML Spy v3.5 NT (http://www.xmlspy.com) by Logan Poelmann -->
<person name="ABCD" address="ABCD" phone="999-999-999-9999" email="ABCD" DOB="1999-12-12" sex="Male | Female | Other"/>
```
simpleTypes

• Many built into the schema spec including: binary, short, long, unsignedShort, string, date, dateTime, timePeriod, ...
• You may define new simpleTypes in your schema or extend the existing types.
• Have the following facets:
  – Enumeration list ex: “FR US UK”
  – minLength
  – maxlength
  – Length
  – Required
  – Optional
  – Default
  – Prohibited
  – Many more …
complexTypes

• You may define new complexTypes in your schema or extend already defined complexTypes types.
• An aggregate datatype i.e. contains multiple values of different datatypes. (Much like a C struct).
• Ex:
  
  ```xml
  <xsd:element name="Person">
    <xsd:complexType>
      <xsd:attribute name="Alive" type="xsd:boolean"/>
      <xsd:attribute name="Age" type="xsd:date"/>
    </xsd:complexType>
  </xsd:element>
  ```

• Have the following facets:
  – Enumeration list ex: “FR US UK”
  – minLength
  – maxlength
  – Length
  – Required
  – Optional
  – Default
  – Prohibited
  – Many more …
Generated XML Schema (.xsd file) for an Invoice
Part of schema file text

<?xml version="1.0" encoding="UTF-8"?>
<!--W3C Schema generated by XML Spy v3.5 NT (http://www.xmlspy.com)-->
<xsd:schema xmlns:xsd="http://www.w3.org/2000/10/XMLSchema" elementFormDefault="qualified">
  <xsd:element name="Address">
    <xsd:complexType>
      <xsd:sequence>
        <xsd:element ref="Dept"/>
        <xsd:element ref="Building"/>
        <xsd:element ref="Street"/>
        <xsd:element ref="City"/>
        <xsd:element ref="State"/>
        <xsd:element ref="Country"/>
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
  <xsd:element name="AreaCode" type="xsd:string"/>
  <xsd:element name="BillTo">
    <xsd:complexType>
      <xsd:sequence>
        <xsd:element ref="PartnerName"/>
        <xsd:element ref="Phone"/>
        <xsd:element ref="Address"/>
      </xsd:sequence>
    </xsd:complexType>
  </xsd:element>
</xsd:schema>

Defines a complexType called an **address**
Defines an element as of simple type **string**
Defines a complexType called an **BillTo** that contains an **address**
Shows an enumeration in the schema that limits the allowable values for an attribute (currency) of an element (Total)
<table>
<thead>
<tr>
<th>Simple Type</th>
<th>Examples (delimited by commas)</th>
</tr>
</thead>
<tbody>
<tr>
<td>string</td>
<td>confirm this is electric</td>
</tr>
<tr>
<td>normalizedString</td>
<td>confirm this is electric</td>
</tr>
<tr>
<td>token</td>
<td>confirm this is electric</td>
</tr>
<tr>
<td>byte</td>
<td>-1, 128</td>
</tr>
<tr>
<td>unsignedByte</td>
<td>0, 126</td>
</tr>
<tr>
<td>base64Binary</td>
<td>GpMF</td>
</tr>
<tr>
<td>hexBinary</td>
<td>OF87</td>
</tr>
<tr>
<td>integer</td>
<td>-125799, -1, 0, 1, 125799</td>
</tr>
<tr>
<td>positiveInteger</td>
<td>1, 125799</td>
</tr>
<tr>
<td>negativeInteger</td>
<td>-125799, -1</td>
</tr>
<tr>
<td>nonNegativeInteger</td>
<td>0, 1, 125799</td>
</tr>
<tr>
<td>nonPositiveInteger</td>
<td>-125799, -1, 0</td>
</tr>
<tr>
<td>int</td>
<td>-1, 12579992674764</td>
</tr>
<tr>
<td>unSignedInt</td>
<td>-1, 125799926743233</td>
</tr>
<tr>
<td>short</td>
<td>-1, 125799926743233</td>
</tr>
<tr>
<td>unSignedShort</td>
<td>0, 1, 125799926743233</td>
</tr>
<tr>
<td>decimal</td>
<td>-1.23, 0, 123.4, 1000.00</td>
</tr>
<tr>
<td>float</td>
<td>-INF, -1E4, -0, 0, 12.78E2, 12, INF, NaN</td>
</tr>
<tr>
<td>double</td>
<td>-INF, -1E4, -0, 0, 12.78E2, 12, INF, NaN</td>
</tr>
<tr>
<td>boolean</td>
<td>true, false</td>
</tr>
<tr>
<td>time</td>
<td>13:20:00.000, 13:20:00.000-05:00</td>
</tr>
<tr>
<td>dateTime</td>
<td>1999-05-31T13:20:00.000-05:00</td>
</tr>
<tr>
<td>duration</td>
<td>P1Y2M3DT1H8M12.38S</td>
</tr>
<tr>
<td>gMonth</td>
<td>39311</td>
</tr>
<tr>
<td>gYear</td>
<td>-06-</td>
</tr>
<tr>
<td>gYearMonth</td>
<td>1999-02</td>
</tr>
<tr>
<td>gMonthDay</td>
<td>1999-02-01</td>
</tr>
<tr>
<td>gDay</td>
<td>-31</td>
</tr>
<tr>
<td>Name</td>
<td>shipTo</td>
</tr>
<tr>
<td>QName</td>
<td>po:USAAddress</td>
</tr>
<tr>
<td>NCName</td>
<td>USAAddress</td>
</tr>
<tr>
<td>anyURI</td>
<td><a href="http://www.example.com">http://www.example.com</a>,</td>
</tr>
<tr>
<td>language</td>
<td><a href="http://www.example.com/doc.htmlMD6">http://www.example.com/doc.htmlMD6</a></td>
</tr>
<tr>
<td>ID</td>
<td>en-GB, en-US, fr</td>
</tr>
<tr>
<td>IDREF</td>
<td></td>
</tr>
<tr>
<td>IDREFS</td>
<td></td>
</tr>
<tr>
<td>ENTITY</td>
<td></td>
</tr>
<tr>
<td>ENTITIES</td>
<td></td>
</tr>
<tr>
<td>NOTATION</td>
<td></td>
</tr>
<tr>
<td>NMTOKEN</td>
<td>US, Brasil</td>
</tr>
<tr>
<td>NMTOKENS</td>
<td>US UK, Bréal Canada Mexique</td>
</tr>
</tbody>
</table>
XML doc and the schema that defines its structure

-- top of schema would be here --

<xsd:element name="Person" type="PersonType"/>
  
  <xsd:complexType name="PersonType">
    <xsd:sequence>
      <xsd:element name="FirstName" type="xsd:string"/>
      <xsd:element name="Middle1Name" type="xsd:string"/>
      <xsd:element name="Middle2Name" type="xsd:string" minOccurs="0"/>
      <xsd:element name="LastName" type="xsd:string"/>
    </xsd:sequence>
  </xsd:complexType>

-- middle of schema would be here --

<xsd:element name="Passengers">
  
  <xsd:complexType>
    <xsd:sequence>
      <xsd:element name="Passenger" type="PersonType" minOccurs="0" maxOccurs="unbounded"/>
    </xsd:sequence>
  </xsd:complexType>
</xsd:element>

-- bottom of schema would be here --

-- Top of XML doc would be here --

<Passengers>
  
  <Passenger>
    <FirstName>John</FirstName>
    <Middle1Name>Q</Middle1Name>
    <Middle2Name>Q</Middle2Name>
    <LastName>Public</LastName>
  </Passenger>
  
  <Passenger>
    <FirstName>Jane</FirstName>
    <Middle1Name>I</Middle1Name>
    <LastName>Doe</LastName>
  </Passenger>

</Passengers>

-- bottom of XML doc would be here --
XML namespaces

• Prevents name collisions when including more than one schema in an XML document.
• Every XML document has a unnamed name space.
• Additional names spaces are declared in the top of the XML document by the following syntax:

```xml
<yourNameSpace :schema
    xmlns:yourNameSpace="http://www.yourwebsite.org/YourSchema"
    elementFormDefault="qualified">
```
## Rough analogies between different data manipulation technologies

<table>
<thead>
<tr>
<th>XML</th>
<th>OOP (Java)</th>
<th>Functional (C)</th>
<th>RDBMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Document</td>
<td>Object Instance with runtime validation of values by mutator methods</td>
<td>Compiled and executing code with Struct and/or local variables with function(s) to set it that has validation code.</td>
<td>Table instance with stored procedure(s) to validate field values (nor integrity checks, though)</td>
</tr>
<tr>
<td>Element</td>
<td>Object Variable may be a finer grained object or a primitive datatype attribute (Int, Short, String, Bool)</td>
<td>Compiled and executing code with Struct and/or local variables with function to set it that has validation code.</td>
<td>Column/field contents and the column name.</td>
</tr>
<tr>
<td>Attribute</td>
<td>Instance of a primitive datatype object attribute</td>
<td>Primitive datatype with function to set it that has validation code.</td>
<td>Column/field contents and the column name.</td>
</tr>
<tr>
<td>Schema</td>
<td>Package or top Level Class Definition and mutator methods that validate</td>
<td>Struct definition and/or variables. A struct made up of other struct definitions. Including function(s) to set it that has validation code.</td>
<td>Table Definition in the DB schema (not the whole DB schema)</td>
</tr>
<tr>
<td>Element Definition</td>
<td>Finer grained class definition with mutator method(s) that validate values before setting the attributes of the object.</td>
<td>Finer grained struct definition with a function to set it that has validation code.</td>
<td>Column definition including datatype ex: BirthDate:Date) with a stored procedure to set the value that has validation code.</td>
</tr>
<tr>
<td>Attribute Definition</td>
<td>Primitive type definition with mutator validation code.</td>
<td>i.e. Int MyInt; Primitive type definition with function that validates content before setting it..</td>
<td>Column definition including datatype ex: BirthDate:Date) with a stored procedure to set the value that has validation code.</td>
</tr>
<tr>
<td>simpleType</td>
<td>Primitive type definition …</td>
<td>Primitive type definition …</td>
<td>Column definition including datatype ex: BirthDate:Date) with a stored procedure to set the value that has validation code.</td>
</tr>
<tr>
<td>complexType</td>
<td>Class def that has multiple attributes that are of different types</td>
<td>A struct that has mixed data types i.e. ints, strings, longs, and/or other struct defs</td>
<td>User defined type?</td>
</tr>
<tr>
<td>XSL file</td>
<td>A transformation and formatting class</td>
<td>A function that converts/formats/embellishes variables to be viewed on screen like printf();</td>
<td>A view of one or more tables</td>
</tr>
</tbody>
</table>
Related standards

• XSL (eXtensible Stylesheet Language)
  – Defines transformation rules for XML documents (XSLT)
  – In particular, defines how an XML document should be translated into HTML for display.

• DOM (Document Object Model) is an object model
  – Allows for the definition of APIs for accessing and manipulating documents.
  – These APIs exist for Java, Jscript, Perl …

• SAX (Simple API to XML)

• XLL = Xlink + XPointer
  – Allows for links between documents or within documents
  – Allows for navigation (like hypertext in HTML)
  – Makes it possible to go beyond the structured document hierarchical model
Related standards (continued)

- XHTML
  - An XML compliant version of HTML
- XSLFO (eXtensible Stylesheet Language – Formatting Objects)
  - A true style sheet language (like CSS)
  - Allows physical display constraints to be expressed by defining and introducing XML presentation elements
- XQL
  - Query language
  - Uses the Xpath pattern matching language
- SVG (Scalable Vector Graphics)
  - Vector graphic language (images, animations…)
  - Adopted by major players in the vector animation industry (Adobe, Macromedia, Microsoft …)
XML definitions

- **DTD**
The "Document Type Definition" is a part of the original XML 1.0 specification that allows a developer, or standards body, to specify what elements and attributes may be used in a particular type of XML document and what their structure and nesting may be. This is also called the content model or schema of an XML document.

If an XML document conforms with the content model defined by a DTD, it is said to be valid with respect to that DTD.

- **XSLT**
The "eXtensible Stylesheet Language Transformation" is a programming language that allows XML documents to be transformed from one schema to another or into entirely different forms, such as HTML pages, WML cards, or PDF files.

- **XPath**
The "XML Path Language" is a language for addressing and querying the content of XML documents.

- **XLink**
The "XML Linking Language" describes hyperlinking in XML documents and extends the hyperlinking concepts of HTML.

- **XPointer**
The "XML Pointer Language" is a companion standard to Xlink and describes mechanisms for addressing particular parts of a document.

- **XML Schema**
The "XML Schema" is an ongoing effort by the W3C to supplant DTDs with a more flexible and powerful system to describe the structure of conforming XML documents, including provisions for defining datatypes.

- **XHTML**
The "Extensible HyperText Markup Language" is the reformulation of HTML 4.0 based upon XML and will soon supplant HTML as the de-facto standard of the Internet.
XML definitions

- **WML**
The "Wireless Markup Language" is used for WAP phone systems to enable a mobile Internet environment and is entirely based on XML - it is described by one particular DTD, which is part of the WML specification.

- **SVG**
  Scalable Vector Graphics. SVG is an XML application used to describe 2D vector graphics, text and raster images. This enables vector graphics to be defined solely in XML.

- **SMIL**
The "Synchronized Media Integration Language" is a XML document type designed to describe multimedia presentations.

- **DOM**
The "Document Object Model" describes how some XML parsers return the information contained in an XML document. The elements of the XML document are described as nodes of a tree that can be traversed by a programmer.

- **SAX**
The "Simple API for XML" provides another programming model used by some parsers, which is based on events instead of a traversable tree.
<XML> Presentation Layer Technologies

XSL, XSLT, XForms
XSL = Extensible Style sheet Language

- XSL is a language enabling the creation of a target document from an XML source document
- The XSL processor uses style sheets to transform XML documents
- Unlike CSS, XSL can serve to transform, filter, and format data
- XSL is based on XPATH (formalized addressing)
- http://www.w3.org/Style/XSL/
XSL for multi-channel distribution (1/2)

• More and more businesses need to integrate new presentation modes
• Appearance of new formats and the need to ensure the continuity of existing ones
  – AudioTel
  – WAP, PDA
  – HTML browsers
  – Network Computer
  – Minitel (in France)
XSL for multi-channel broadcasting (2/2)

- Separating data from presentation with an XSL Formatting Object

Diagram:

- XML Document
  - XSL Processor
  - WML Style Sheet
  - HTML, VoxML Style Sheet
  - Postscript Style Sheet
  - XML Style Sheet

- WAP Terminal
- HTML browser
- PDF Document
- Specific XML Document
- Partner B2B site
XSL

• Requires that your browser have an XSL engine built in. For IE is msmxml.dll
• Microsoft® XML Parser (MSXML), available from
• Microsoft Internet Explorer 5.0 or later.
• Excellent quick tutorial
<?xml version="1.0" ?>
<sales>
  <summary>
    <heading>Greenwich Village Book Publishing</heading>
    <subhead>Regional Sales Report</subhead>
    <description>Sales report for the West Coast, Central and East Coast regions.</description>
  </summary>
  <data>
    <region>
      <name>West Coast</name>
      <quarter number="1" books_sold="24000" />
      <quarter number="2" books_sold="38600" />
      <quarter number="3" books_sold="44030" />
      <quarter number="4" books_sold="21000" />
    </region>
    <region>
      <name>Central</name>
      <quarter number="1" books_sold="11000" />
      <quarter number="2" books_sold="16080" />
      <quarter number="3" books_sold="25000" />
      <quarter number="4" books_sold="29000" />
    </region>
    <region>
      <name>East Coast</name>
      <quarter number="1" books_sold="27000" />
      <quarter number="2" books_sold="31400" />
      <quarter number="3" books_sold="40100" />
      <quarter number="4" books_sold="30000" />
    </region>
  </data>
</sales>
<?xml version="1.0"?>
<sales>
  <summary>
    <heading>Greenwich Village Book Publishing</heading>
    <subhead>Regional Sales Report</subhead>
    <description>Sales report for the West Coast, Central and East Coast regions.</description>
  </summary>
  <data>
    <region>West Coast</region>
    <quarter number="1">books_sold="24000"</quarter>
    <quarter number="2">books_sold="88600"</quarter>
    <quarter number="3">books_sold="44030"</quarter>
    <quarter number="4">books_sold="21000"</quarter>
    <region>Central</region>
    <quarter number="1">books_sold="11000"</quarter>
    <quarter number="2">books_sold="16080"</quarter>
    <quarter number="3">books_sold="25000"</quarter>
    <quarter number="4">books_sold="29000"</quarter>
    <region>East Coast</region>
    <quarter number="1">books_sold="27000"</quarter>
    <quarter number="2">books_sold="31400"</quarter>
    <quarter number="3">books_sold="40100"</quarter>
    <quarter number="4">books_sold="30000"</quarter>
  </data>
</sales>
Adding an XSL statement

```xml
<?xml version="1.0" ?>
<?xml-stylesheet type="text/xsl" href="Transform.xsl"?>

<sales>
  <summary>
    <heading>Greenwich Village Book Publishing</heading>
    <subhead>Regional Sales Report</subhead>
    <description>Sales report for the West Coast, Central and East Coast regions.</description>
  </summary>
  <data>
    <region>
      ...
    </region>
  </data>
</sales>

//rest of file would be here
```
<?xml version="1.0"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform"
    version="1.0">
    <xsl:output method="html"/>
    <xsl:template match="/>
        <HTML>
            <HEAD>
                <TITLE><xsl:value-of select="//summary/heading"/></TITLE>
            </HEAD>
            <BODY>
                <h1><xsl:value-of select="//summary/heading"/></h1>
                <h2><xsl:value-of select="//summary/subhead"/></h2>
                <p><xsl:value-of select="//summary/description"/></p>
            </BODY>
        </HTML>
    </xsl:template>
</xsl:stylesheet>
Greenwich Village Book Publishing

Regional Sales Report

Sales report for the West Coast, Central and East Coast regions.
Greenwich Village Book Publishing

Regional Sales Report

Sales report for the West Coast, Central and East Coast regions.

<table>
<thead>
<tr>
<th>Region/Quarter</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>West Coast</td>
<td>24,000</td>
<td>38,600</td>
<td>44,030</td>
<td>21,000</td>
<td></td>
</tr>
<tr>
<td>Central</td>
<td>11,000</td>
<td>16,080</td>
<td>25,000</td>
<td>29,000</td>
<td></td>
</tr>
<tr>
<td>East Coast</td>
<td>27,000</td>
<td>31,400</td>
<td>40,100</td>
<td>30,000</td>
<td></td>
</tr>
</tbody>
</table>
Greenwich Village Book Publishing

Regional Sales Report

Sales report for the West Coast, Central and East Coast regions.

<table>
<thead>
<tr>
<th>Region</th>
<th>Quarter</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>West</td>
<td>Coast</td>
<td>24,000</td>
<td>38,600</td>
<td>44,030</td>
<td>21,000</td>
<td>127,630</td>
</tr>
<tr>
<td>Central</td>
<td></td>
<td>11,000</td>
<td>16,080</td>
<td>25,000</td>
<td>29,000</td>
<td>81,080</td>
</tr>
<tr>
<td>East</td>
<td>Coast</td>
<td>27,000</td>
<td>31,400</td>
<td>40,100</td>
<td>30,000</td>
<td>128,500</td>
</tr>
</tbody>
</table>
<?xml version="1.0"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform" version="1.0">
<xsl:output method="html"/>
<xsl:template match="/">

<HTML>
<HEAD>
<TITLE><xsl:value-of select="//summary/heading"/></TITLE>
</HEAD>
<BODY>
<h1><xsl:value-of select="//summary/heading"/></h1>
<h2><xsl:value-of select="//summary/subhead"/></h2>
<p><xsl:value-of select="//summary/description"/></p>

<!--Part 2-->
<table>
<tr>
<th>Region\Quarter</th>
<xsl:for-each select="//data/region[1]/quarter">
<th>Q<xsl:value-of select="format-number(@books sold,'###.###')"/></th>
</xsl:for-each>
</tr>
<tr>
<th>Total</th>
</tr>
<xsl:for-each select="//data/region">
<tr>
<th style="text-align:left"><xsl:value-of select="name"/></th>
<xsl:for-each select="quarter">
<td style="text-align:right"><xsl:value-of select="format-number(@books sold,'###.###')"/></td>
</xsl:for-each>
</tr>
</xsl:for-each>
</table>
</BODY>
</HTML>
XPath

• “XPath - is a language for addressing parts of an XML document, designed to be used by both XSLT and XPointer. In addition to its use for addressing, XPath is also designed so that it has a natural subset that can be used for matching (testing whether or not a node matches a pattern); this use of XPath is described in XSLT.

XPath models an XML document as a tree of nodes. There are different types of nodes, including element nodes, attribute nodes and text nodes. XPath defines a way to compute a string-value for each type of node. "

- W3C
XPath Examples

- `.author`
  
  Find all author elements within the current context (both are equivalent).

- `first.name`

  Find all first.name elements.

- `/bookstore`

  Find the bookstore element at the root of this document.

- `/*`

  Find the root element of this document.

- `//author`

  Find all author elements anywhere within the current document.

- `/bookstore[@specialty = "textbooks"]`

  Find all bookstores where the value of the `specialty` attribute is equal to "textbooks".

- `book[@style = /bookstore/@specialty]`

  Find all books where the value of the `style` attribute on the book is equal to the value of the `specialty` attribute of the bookstore element at the root of the document.

- `book[@style = /bookstore/@specialty]`

  Find all books where the value of the `style` attribute on the book is equal to the value of the `specialty` attribute of the bookstore element at the root of the document.
XML to HTML via an XSL

<?xml version="1.0" encoding="UTF-8"?>
<!-- edited with XML Spy v4.0 U (http://www.xmlspy.com) by Logan Poelman (Valtech) -->

<documentid>994-2001-123x</documentid>
<manufacturer>Altova</manufacturer>
<website>http://www.altova.com</website>
<productline>123xml</productline>
<security>
  <classification>confidential</classification>
  <level>2</level>
  <effective>2001-03-26</effective>
  <use internal="true" external="false"/>
</security>
<tracking>
  <modification operation="created">
    <date>2001-01-15</date>
    <time>17:45:09</time>
    <by>Ben Abell</by>
    <department>IT Admin</department>
    <status>in progress</status>
    <changes>created from previous 123-400 datasheet</changes>
  </modification>
  <modification operation="edited">
    <date>2001-02-20</date>
    <time>09:27:12</time>
    <by>Tina Eisinger</by>
    <department>Marketing</department>
    <status>in progress</status>
    <changes>highlighting new feature sets</changes>
  </modification>
</tracking>
Sample XSL style sheet

<?xml version="1.0" encoding="UTF-8"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform" version="1.0">
  <xsl:output method="xml" version="1.0" encoding="UTF-8" omit-xml-declaration="no" indent="no" media-type="text/html"/>
  <!--XSL Stylesheet for generating Datasheet-->
  <xsl:template match="datasheet">
    <html>
      <head>
        <title>
          <xsl:value-of select="manufacturer"/>
          <xsl:value-of select="productline"/>
        </title>
      </head>
      <body>
        <p align="right">
          <img src="Altova_right_300.gif" alt="Altova logo"/>
        </p>
        <h1>
          <xsl:value-of select="manufacturer"/>
          <xsl:value-of select="productline"/>
        </h1>
      </body>
    </html>
  </xsl:template>
</xsl:stylesheet>

- - - rest of XSL would be here - - -
XSLT

• XML Stylesheet Language Transformation
• Subset of XSL.
• Designed for transforming one XML doc to a different format.
• Specialty language targeted at transformations used to format documents for presentation including: HTML, WML, PDF, etc.
• An XSLT file is an XML file.
• Patterns match source tree (source XML file) are transformed and appended to the output tree (target XML file).
XForms

- XForms are comprised of separate sections that describe what the form does, and how the form looks. This allows for flexible presentation options, including classic XHTML forms, to be attached to an XML form definition.

- **Key Goals of XForms**
  - Support for handheld, television, and desktop browsers, plus printers and scanners
  - Richer user interface to meet the needs of business, consumer and device control applications
  - Decoupled data, logic and presentation
  - Improved internationalization
  - Support for structured form data
  - Advanced forms logic
  - Multiple forms per page, and pages per form
  - Suspend and Resume support
  - Seamless integration with other XML tag sets

- [http://www.w3.org/TR/xforms/](http://www.w3.org/TR/xforms/)
Sample XForm

The xform is MVC it has a model section, a View (Data via a CSS / XSL) and Controller (action)

```html
<html xmlns="http://www.w3.org/1999/xhtml"
 xmlns:xforms="http://www.w3.org/2002/01/xforms"
 xmlns:ev="http://www.w3.org/2001/xml-events"
 xmlns:xlink="http://www.w3.org/1999/xlink">
<head>
  <title>XForms: Order Form</title>
  <link rel="stylesheet" type="text/css" href="AES.css">
</head>

<form>
  <table>
    <tr><th>First Name</th><th>Initial</th><th>Last Name</th></tr>
    <tr><td>Joel</td><td>1</td><td>Rivera-Gonzalez</td></tr>
    <tr><th>Street Address</th></tr>
    <tr><td>253 My Street</td></tr>
    <tr><th>City</th><th>State</th><th>Zip Code</th></tr>
    <tr><td>This Town</td><td>NY</td><td>11122</td></tr>
    <tr><th>Sub Total</th><td>$ 22.45</td></tr>
    <tr><th>Tax Amount</th><td>$ 1.91</td></tr>
    <tr><th>Total</th><td>$ 24.36</td></tr>
  </table>
</form>

Shopping Cart

<table>
<thead>
<tr>
<th>Quantity</th>
<th>Item Description</th>
<th>Price p/u ($)</th>
<th>Total ($)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>White Chocolate Bars</td>
<td>1.45</td>
<td>7.25</td>
</tr>
<tr>
<td>8</td>
<td>Dark Chocolate Bar</td>
<td>1.45</td>
<td>11.60</td>
</tr>
<tr>
<td>2</td>
<td>Caramel Filled chocolate</td>
<td>1.80</td>
<td>3.60</td>
</tr>
</tbody>
</table>
```

Submit  Cancel
XForms – model fragment

```xml
<xforms:model id="data">
  <xforms:instance>
    <OrderInfo>
      <PersonalInfo>
        <Name>
          <First></First>
          <Middle></Middle>
          <Last></Last>
        </Name>
        <Address>
          <Street></Street>
          <City></City>
          <State></State>
          <Zip></Zip>
        </Address>
      </PersonalInfo>
      <PriceInfo>
        <SubTotal></SubTotal>
        <TaxTotal></TaxTotal>
        <TaxRate></TaxRate>
        <Total></Total>
      </PriceInfo>
      <TaxInfo>
        <CT>.060</CT>                  <NY>.085</NY>  <NJ>.083</NJ>
      </TaxInfo>
      <ShoppingCart>
        <ProductInfo name="itm1">
          <Quantity>5</Quantity>
          <Description>Wht. Chocolate Bars</Description>
          <UnitPrice>1.45</UnitPrice>
          <ItemTotal>7.25</ItemTotal>
        </ProductInfo>
      </ShoppingCart>
    </OrderInfo>
  </xforms:instance>
</xforms:model>
```
XForms – action (controller) fragment

```xml
<body>
  <xforms:group>
    <div>
      <xforms:input class="First" id="first"
                    ref="OrderInfo/PersonalInfo/Name/First">
        <div>
          <xforms:label>First Name</xforms:label>
        </div>
        <xforms:hint>Enter First Name</xforms:hint>
      </xforms:input>
      <xforms:input class="Middle" id="middle"
                    ref="OrderInfo/PersonalInfo/Name/Middle">
        <xforms:label class="label">Middle Initial</xforms:label>
        <xforms:hint>Enter Middle Initial</xforms:hint>
      </xforms:input>
      <xforms:input class="Last" id="last"
                    ref="OrderInfo/PersonalInfo/Name/Last">
        <xforms:label class="label">Last Name</xforms:label>
        <xforms:hint>Enter Last Name</xforms:hint>
      </xforms:input>
    </div>
  </xforms:group>
</body>
```
XQL - XML Query Language

“The XML Query Language (XQL) is a notation for addressing and filtering the elements and text of XML documents. XQL is a natural extension to the XSL pattern syntax. It provides a concise, understandable notation for pointing to specific elements and for searching for nodes with particular characteristics.

The XSL pattern language provides an extremely understandable way to describe a class of nodes to process. It is declarative rather than procedural. One simply describes the types of nodes to look for using a simple pattern modeled after directory notation. For example, book/author means find author elements contained in book elements.

XQL provides a natural extension to the XSL pattern language. It builds upon the capabilities XSL provides for identifying classes of nodes, by adding Boolean logic, filters, indexing into collections of nodes, and more.

XQL is designed specifically for XML documents. It is a general purpose query language, providing a single syntax that can be used for queries, addressing, and patterns. XQL is concise, simple, and powerful.”

- W3C
Managing `<XML>` Persistence
Relational Theory

Tables can be joined through the key (ID) to create a one to many relationship.

Customer Table

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>0012</td>
<td>Poelman</td>
<td>43 Mercer</td>
</tr>
<tr>
<td>0013</td>
<td>Smith</td>
<td>2000 Main</td>
</tr>
<tr>
<td>0324</td>
<td>Chang</td>
<td>55 Broadway</td>
</tr>
</tbody>
</table>

Invoice Table

<table>
<thead>
<tr>
<th>ID</th>
<th>InvoiceNum</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0012</td>
<td>12324</td>
<td>344.55</td>
</tr>
<tr>
<td>0012</td>
<td>44424</td>
<td>120.00</td>
</tr>
<tr>
<td>0324</td>
<td>55554</td>
<td>2344.00</td>
</tr>
<tr>
<td>0013</td>
<td>43533</td>
<td>0.99</td>
</tr>
</tbody>
</table>
A Join between both tables where ID = 0012 yields the table

<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Address</th>
<th>InvoiceNum</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0012</td>
<td>Poelman</td>
<td>43 Mercer</td>
<td>12324</td>
<td>344.55</td>
</tr>
<tr>
<td>0012</td>
<td>Poelman</td>
<td>43 Mercer</td>
<td>44424</td>
<td>120.00</td>
</tr>
</tbody>
</table>
As XML it could be represented as:

```xml
<CustomerInvoices>
  <Customer>
    <ID>0012</ID>
    <Name>Poelman</Name>
    <Address>43 Mercer</Address>
    <Invoice>
      <Number>12324</Number>
      <Total>344.55</Total>
    </Invoice>
    <Invoice>
      <Number>44424</Number>
      <Total>120.00</Total>
    </Invoice>
  </Customer>
  <Customer>
    <ID>0013</ID>
    <Name>Smith</Name>
    <Address>2000 Main</Address>
    <Invoice>
      <Number>43533</Number>
      <Total>0.99</Total>
    </Invoice>
  </Customer>
</CustomerInvoices>
```
XML DataBase Management System (XML DBMS / XDBMS)

- XML is NOT relational its hierarchical!
- Storage of XML is inefficient in:
  - Files
  - RDBMS
  - Object DB (maybe but XML is just data no methods)
- XML database
  - XML data is parsed and accessible through a DOM-like API and SQL-type queries
  - High-performance, multi-user processing
  - Stored in a form that is more efficient for hierarchical data
- XML front end permitting access to heterogeneous data, relational data in particular
- Supports XQL
XML DBMS

- 4Suite, 4Suite Server
- Birdstep RDM XML
- Centor Interaction Server
- Cerisent XQE
- Coherity XML Database
- DBDOM
- dbXML
- DOM-Safe
- eXist
- eXtc
- eXtensible Information Server (XIS)
- GoXML DB
- Infonyte DB
- Ipedo XML Database
- Lore
- Lucid XML Data Manager
- MindSuite XDB

- Natix
- Neocore XML Management System
- Ozone
- Socrates XML
- Tamino
- Tendara Mobile XML Database
- TeraText DBS
- TEXTML Server
- Virtuoso
- XDBM
- XDB
- Xfinity Server
- X-Hive/DB
- Xindice
- Xylene Zone Server
- XYZFind Server

Good overview http://www.rpbourret.com/xml/XMLAndDatabases.htm
Using `<XML>` in the enterprise
Integration of company data

- XML database as a front end:
  - Data durability and accessibility
  - Architecture independence
XML for application integration

- Electronic business must confront the growing number of information systems products (ERP, Business Directories, SGBD …)
- EAI (Enterprise Application Integration) offers a unified vision of heterogeneous data sources
- Modeling business processes in a standard way to mask the complexity of technical implementations
  - As a workflow graph
    - Works by sending asynchronous (or synchronous) messages
- Workflow interacts with products through event-linked connectors (new order, etc.)

The products must provide some EAI connectors!
XML for application integration (EAI)
Protocols, technologies and such
BEEP

• Blocks Extensible Exchange Protocol
  – Connections
  – Authentication
  – packaging up at the TCP/IP level of the messages

• SOAP deliberately avoids these

• BEEP really competes on the same level as HTTP.
BEEP

• HTTP
  – Stateless
  – request-response-oriented protocol
  – No requests with a response
  – No response without a request
  – No state is preserved between requests
  – Client/Server not Peer to Peer
  – Synchronous not asynchronous
BEEP

• BEEP
  – Stateful
  – Connection oriented
    • Session is the connection
    • Ordered, reliable and congestion sensitive.
  – Peer oriented or client/server
  – Message oriented protocol
  – Asynchronous
  – Pseudo synchronous is possible – send message and block waiting for a response.
  – Multiplexed - multiple channels communicating over a single transport stream
BEEP - concerns

• Framing
• Encoding
• Reporting
• Multiplexing
• Authentication
• Privacy
• Dev language independent – Java, C++, C, Python, Ruby, …
BEEP

- MIME encoded payloads
- Uses SASL & TLS (SSL)
- peer that starts a connection as the *initiator*, and the peer accepting the connection as the *listener*. When a connection is established between the two, a BEEP session is created.
- BEEP is a framework for building application protocols. Like XML is for docs, BEEP is for protocols.
- [http://www.beepcore.org/beepcore/home.jsp](http://www.beepcore.org/beepcore/home.jsp)
DIME

- Direct Internet Message Encapsulation
- DIME is a message format not a protocol
- DIME is a *binary* format: fields have fixed size
  
  “The design goal of DIME is to provide an efficient and simple message format that can accommodate the following: (1) Encapsulating arbitrary documents and entities, including encrypted data, XML documents, XML fragments, image data like GIF and JPEG files, etc. (2) Encapsulating documents and entities initially of unknown size. This capability can be used to encapsulate dynamically generated content or very large entities as a series of chunks. (3) Aggregating multiple documents and entities that are logically associated in some manner into a single message.”
Jabber

- XMPP - Extensible Messaging and Presence Protocol
- Client / Server based Instant Msging – end users talk through a servers communicating with each other.
- Like “instant email” but connection oriented. It does queue messages for you if you are are disconnected.
- Jabber addresses are of the same form as email addresses
- Can be used between applications to pass data
- XML based
- TCP socket over port 5222
- http://www.jabber.org/
Portals — technology that aggregates together multiple web interfaces into a single browser view.

*if viewing one website, is good many must be better*
Portal

• portal

1. A doorway, entrance, or gate, especially one that is large and imposing.
2. An entrance or a means of entrance: the local library, a portal of knowledge.
3. A website considered as an entry point to other websites, often by being or providing access to a search engine.

- www.dictionary.com
Analyst View on Portal Value

"The Portal is one response to several integration problems that exist in corporate IT... The need for enterprises to integrate disparate application systems, knowledge and information sources is acute and growing."

- Ovum Evaluates: Portal Software
  Ovum Group
  Nov 2001
Analysts Define a Portal

- A browser page that aggregates together two or more web pages in a single view
  - Simplest definition

- It probably would be a good idea to start off with a definition. In a portal, we are looking at the aggregation of unstructured content — things like documents and threaded discussion databases — as well as structured data (data warehousing type data, and IT services, business intelligence or collaboration) all kind of rolled up into a business context
  - META Group

- A portal provides personalised access to an appropriate range of information for a particular purpose
  - Ovum

- A personalized Web site built on a common set of application services that enables firms to connect employees, customers, partners, and suppliers to all appropriate corporate computing resources from any Internet-capable device
  - Forrester
Portal “desktop” view

- Market Summary
- Address Book
- Weather
- Bill Pay
- Calculator
- My Front Page Headlines
- Top Stories from Reuters
- Sports News from Reuters
- Entertainment News from Reuters
- Notepad
- Movie Releases

Portlet
Define a Portal

Other features that may/may not define a portal:

- **Customization** - user can change content sources (web pages & applications) and layout of their personal view.

- **Searching**
  
  Federated Search - search across multiple databases.
  
  Mixed content type search - images, structured, semistructured and structured data.
  
  Meta data enabled search.

- **Personalization** - site changes content and layout to match rules set up by administrator. These rules can take into account the specific user’s preferences, previous behavior, date, special offers, etc.

- **Portlet to Portlet communication** - portlets are the parts of the portal view that contains a single content source/application.

- **Standardized but changeable “look & feel”**
Particular “parts” of a portal desktop

- **Desktop Tabs**
- **Portlets**
- **Portlet Execution Mode**
- **Portal Titlebar**
- **Portlet Titlebar**
Portal Engine Web-enablement

Connect to all appropriate corporate computing resources from any Internet-capable device

- Data Warehouse & Business Intelligence
- Content & Site Cataloging (e.g., search)
- Community & Member Interactions
- Collaboration & Communication
- Line-of-Business Apps & Operational Packages
- Syndicated Services (e.g., news feeds)
Example Portal User Interface
Case Kmart

- **Functionalities**
  - electronic paycheck presentment
  - W-2 statement processing services from PeopleSoft
  - online tutorials for software applications
  - sales reports
  - key performance metrics
  - e-mail
  - a global personnel directory
  - retail stock prices
  - a Kmart store locator

- **Benefits**
  - improves customer service, speeding searches for information
  - drives growth, delivering sales and marketing tools to the desktop
  - increases employee satisfaction, speeding HR processes online
Portal User Communities

- **B2C** - Business to Customer
- **B2E** - Business to Employee
- **B2B** - Business to Business
- **B2CSR2C** - Business to CSR to Customer
- **G2C** - Government to Citizen
- **C2C** - Customer to Customer
WebSphere Portal Server
Plumtree Portal Server
A better Java than Java!
.NET According to Microsoft

"Quite simply, .NET is Microsoft's platform for XML Web services. XML Web services allow applications to communicate and share data over the Internet, regardless of operating system or programming language.

The Microsoft® .NET Platform includes a comprehensive family of products, built on XML and Internet industry standards, that provide for each aspect of developing, managing, using, and experiencing XML Web services. XML Web services will become part of the Microsoft applications, tools, and servers you already use today and will be built into new products to meet all of your business needs."

- Microsoft website September 25, 2001
.NET Distilled - What

- **OO layer that runs on top of the OS**
  - Windows 9x, 2000, XP, PocketPC, Linux?
  - Simplify distributed object apps
  - Gives a path to migrate
  - CLR is a Virtual Machine (VM) + Libraries
  - Managed code model

- **Language independent**

- **Better, Web-friendly ASP.NET architecture**
  - XCOPY deployment
  - Lazy instantiation

- **New security**

- **New UI model**
.NET Framework

Common Language Specification (CLS)

ADO.NET: Data and XML

ASP.NET: Web Services & Web Forms

Windows Forms

Common Type System (CTS)

Common Language Runtime (CLR)

Operating Systems (9x, ME, 2000, XP, PocketPC, Linux?)

Visual Studio.NET
.NET Distilled - How

- Framework and tools
- Distributed application platform based on mostly industry standards
- A set of languages
- A common inheritance method across languages (CTS)
- Shared classes (CLS)
- Common exception model across languages
- Server products
- VisualStudio.NET tool
- Wizards/classes for the average developer
.NET Distilled - Why

- Technology leadership mindshare
- A better mousetrap
- Interoperability
- Potential instability of poorly written COM components
- Windows wasn’t OO
- Java – language, platform, ideas
- DNA / ASP requirement multiple technologies
- Piracy of Microsoft software
Microsoft and 3\textsuperscript{rd} party support for .NET development includes (announced):

- APL
- C++
- C#
- COBOL
- Component Pascal
- Curriculum
- Eiffel
- Fortran
- Haskell
- Java Language
- Microsoft JScript
- Mercury
- Mondrian
- Oberon
- Oz
- Pascal
- Perl
- Python
- RPG
- Scheme
- Smalltalk
- Standard ML
- Microsoft Visual Basic
Vision

• Platform Independence
• Common Libraries Across Languages
• Embracing Open Standards
• No Need For Java
• Software As Services
• Enterprise Software Infrastructure
• .NET Servers - Off The Shelf Solutions
• Lower Time To Market
• Lower TCO
Benefits

• Enabling technology – what CORBA / COM+ / RMI promised may now be delivered
• Faster time to market
• Lower development costs
• Buy vs. Build
• Lower footprint so lower TCO
• Integration between packaged software (MS OfficeXP) and custom apps
Benefits

• OO
• Distributed
• Enterprise scalable
• Better security
• Monitorable
• More “plumbing” is provided by the OS
• Simpler integration - A2A, EAI, B2B
• Lower skills needed by developers
C#.NET Upside

- **Objected-Oriented** – Encapsulation, Polymorphism, Inheritance
- **Versioning Support** – Eliminate DLL hell to prevent code incompatibility when newer code are made.
- **Type Safety** – All variable casts are required to be safe. Can’t create an invalid cast.
- **Rich API** - Pre-built libraries allows programmer to be more productive.
- **Leverage the CLR environment** – Automatic Garbage Collection, Security, Language Interoperability, and Legacy Integration.
C# .NET Downside

- **Lack of Portability** - Only supported on the Window platform. It is currently not portable across other operating systems.
- **New syntax and API** - Applying the new class libraries and constructs effectively may require a slight learning curve.
- **YALL** - Yet another language to learn and support.
- **Unproven** - no track record or its ease of use, power, productivity,
C# and Java Key Similarities

• Code compiled to interpreted language (IL/byte-code).
• Automatic garbage collection of objects.
• All values are initialized before use.
• Interfaces, with multiple inheritance of interfaces, single inheritance of implementation.
• All code scoped to packages (or assemblies in .NET)
Creating a WebForms app in .net
Design “canvas”

Click and it generates a handler in C#
using System;
using System.Collections;
using System.ComponentModel;
using System.Data;
using System.Drawing;
using System.Web;
using System.Web.SessionState;
using System.Web.UI;
using System.Web.UI.WebControls;
using System.Web.UI.HtmlControls;

namespace WebApplication1
{
    /// <summary>
    /// Summary description for WebForm1.
    /// </summary>
    public class WebForm1 : System.Web.UI.Page
    {
        protected System.Web.UI.WebControls.Label Label1;
        protected System.Web.UI.WebControls.TextBox txtYourName;
        protected System.Web.UI.WebControls.TextBox txtTargetArea;

        private void Page_Load(object sender, System.EventArgs e)
        {
            // Put user code to initialize the page here
        }

        #region Web Form Designer generated code
        override protected void OnInit(EventArgs e)
        {
            // CODEGEN: This call is required by the ASP.NET Web Form Designer.
            InitializeComponent();
            base.OnInit(e);
        }

        /// <summary>
        /// Required method for Designer support - do not modify
        /// the contents of this method with the code editor.
        /// </summary>
        private void InitializeComponent()
        {
            this.btnMoveText.Click += new System.EventHandler(this.btnMoveText_Click);
            this.Load += new System.EventHandler(this.Page_Load);
        }
        #endregion

        private void btnMoveText_Click(object sender, System.EventArgs e)
        {
        }
    }
}
Add one line of code

- `txtTargetArea.Text = txtYourName.Text;`
- And you have a web application!
.NET Recap

• Great New Development Languages and IDE
• Extremely Powerful VM and Container Model (CLR)
• Language Interoperability A Big Plus
• XML everywhere!
• Fast Dev for Small - Medium Web/Desktop Apps
• Not Enterprise Scale Yet but will be soon
• Better Security
• Mixed COM+/.NET highly likely on next 12-24 months
• TCO could be higher over next 24 months but lower in a future CLR / OS without COM+
## .NET vs. J2EE

<table>
<thead>
<tr>
<th></th>
<th>.NET</th>
<th>J2EE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OO Languages</strong></td>
<td>C#, VB, …</td>
<td>Java</td>
</tr>
<tr>
<td><strong>Virtual Machine</strong></td>
<td>CLR</td>
<td>JVM</td>
</tr>
<tr>
<td><strong>Container</strong></td>
<td>ServicedComponents in the COM+</td>
<td>EJB</td>
</tr>
<tr>
<td><strong>Pseudo ops</strong></td>
<td>IL</td>
<td>JavaByteCode</td>
</tr>
<tr>
<td><strong>Platforms</strong></td>
<td>Windows</td>
<td>Windows, Unix, OSS/390</td>
</tr>
<tr>
<td></td>
<td>Mono on Linux</td>
<td></td>
</tr>
<tr>
<td><strong>XML</strong></td>
<td>Everywhere</td>
<td>Lots</td>
</tr>
<tr>
<td><strong>WebServices</strong></td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>UI</strong></td>
<td>WinForms, WebForms</td>
<td>Swing, JSP</td>
</tr>
<tr>
<td><strong>Maturity</strong></td>
<td>Beta but reused stuff and good quality now</td>
<td>3-4 years</td>
</tr>
</tbody>
</table>

Copyright Poelman & Associates, Inc. 2002 - All Rights Reserved
Distributed Computing Technologies

What fits best where?
Interface Domains

- B2B – Business to Business
- L2L – Line of Business to Line of Business
- S2S – System to System
- A2A - Application to Application
- Application
Aspects of Interface Domains

- $$ Value of Operation
- Trust
- Bandwidth Available
- Platform Sensitivity
- Administrative “ownership”
- Functionality Level
- Processactional
- Transactional
- Functional
Interface Protocol Has …

• Lookup Service
  – Get the “Address” of the other party
• Handshaking
• Information Format
• Interaction Management
• Security Support
• Processaaction Support
• Exception Support
• Communications Model - Blocking vs. Non-blocking
  – (A.K.A. Synchronous vs. non-synchronous)
• Paradigm
  – Call / Return, Request / Reply, Pub / Sub, Event Handling
• Wire Transport
Interface Protocols

- WS – Web Services
- HTTP v1.1
- Messaging – MQ Series, MSMQ, JMS
- SMTP / POP3 / IMAP - email
- .NET Object Remoting via HTTP
- .NET Object Remoting via TCP
- CORBA – Common Object Request Broker Architecture
- RMI – Java Remote Method Invocation
- COM/DCOM – Distributed Component Object Model
- Sockets – Low-level TCP/IP
- File Sharing / exchange - earliest example of EAI
- Local Function / Method Invocation - standard calls
Interface Protocol Has …

• A Purpose or Goal!!
  – **WS** - provide a firewall friendly, language neutral RPC that supported looking up services in a directory. A SOA over the Internet.
  – **HTTP** – Initially to share hypertext based documents over the IP based internet. HTML based documents. Many clients, few servers, scalable.
  – **.NET Object Remoting via HTTP** - like DCOM that works over the Internet
  – **.NET Object Remoting via TCP** - like DCOM that is easier and more flexible over your LAN.
  – **RMI** – simple way to allow remote invocation of Java objects over IP network.
  – **IIOP** – allow different CORBA ORBs to interoperate
  – **CORBA** – distributed object applications. OOP. C++ oriented.
  – **MQ (Messaging Queuing)** – platform independent, language independent, time independent, A2A integration.
  – **COM/DCOM** - allow remotely calling functionality on one or more windows boxes on a LAN.
Interface Domain Facets

- Application
  - A2A
  - S2S
  - L2L
  - B2B

- Administrative
  - Control

- Low Bandwidth
  - Available

- Platform Neutrality
  - Implicit Trust Relationship

- Processional
  - Transactional

- High Volume
  - High Value
  - High Level Functionality
Protocol Applicability

TCP/IP Sockets
Message Queuing
.COM/DCOM/COM+
TCP/IP Sockets
CORBA
.NET HTTP Remoting
.NET TCP Remoting
RMI
.NET TCP Remoting
CORBA
Message Queuing
TCP/IP Sockets

Application
A2A
S2S
L2L
B2B
Web Service
User Interface

GUI

Application

WS - Messaging

WS - CORBA - Messaging

CORBA - COM+  .NET HTTP Remoting - Messaging

RMI - CORBA - COM+  .NET TCP Remoting - Messaging

RMI - CORBA - COM+  .NET TCP Remoting Local Func / Method Invocation

Copyright Poelman & Associates, Inc. 2002 - All Rights Reserved
The end!
Find out more...

- Interesting sites:
  - http://w3c.org
  - http://www.w3.org/TR/
  - http://www.oasis-open.org
  - http://www.schema.net
XML web links

- http://www.w3.org/TR/SOAP/
- http://www.xml.org
- http://www.UDDI.org
- http://xml.com/xml/pub
- http://java.sun.com/xml
- http://www.xmlsoftware.com
- http://www.xml.org/xml/resources_cover.shtml
• http://www.microsoft.com/net/
• Interesting tutorials http://www.w3schools.com
• http://www.gotdotnet.com/