What is XML
and
Why do you care ?

Ellen McCubbin
ellenmcc@us.ibm.com

Trademark information

- UNIX is a registered trademark in the United States, other countries, or both and is licensed exclusively through X/Open Company Limited.
- Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.
- Microsoft, Windows, Windows NT, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.
- IBM is a registered trademark of the IBM Corporation in the United States, other countries, or both.

Agenda

- Why XML ?
- XML in General
- XML versus HTML
- How will XML be Used ?
- What are the challenges in using XML?
- XML Standards Bodies
- Retail XML Initiatives
- Sources of More information

The Web Has Changed Our Expectations

- Everything must be connected
- Information must be searchable
- Data one universal format
- Data seamlessly presented
- Documents must be dynamically assembled from smaller pieces
- What is the difference between “structured” and “unstructured” data ?

QUIZ

Quiz - what does the acronym GML stand for ?

HTML

- Defines how data should look
- Fixed Set of element types
- Designed for the web
- Documents have a long life span

However, todays web applications need Enterprise Wide Integration, across supply chains. HTML does not provide this capability.

HTML - just for formatting

```
<HTML>
<HEAD><TITLE>E-Invoice</TITLE></HEAD>
<BODY>
<H1>Automobile E-Invoice</H1>
<H2>Purchaser:</H2>
Eric Severson, Executive Consultant<BR>
IBM Corporation
<H2>Item Purchased:</H2>
1997 Black Jeep Wrangler<BR>
Purchase Date: 7/1/96<BR>
Purchase Price: $20,000
</BODY></HTML>
```

XML - a short history

- Simplified subset of SGML
- Effort started in 1996
- Official Standards Body - W3C
- Provides for "Smart Data"
- Terminology borrowed from Data Modeling Disciple
  - Entities
  - Attributes

XML Design Goals - an excerpt

- "Straighforwardly" usable over the Internet
- Support a wide variety of applications
- Compatible with SGML
- Easy to write programs which process XML documents
- Number of optional features to be kept to the absolute minimum, ideally zero
- XML Documents should be human-legible and reasonably clear
- Design of XML shall be formal and concise
- XML Documents shall be easy to create

XML

- Defines what data means
- Create your own tags
- Middle Tier server required
- Designed for internet technologies
- Documents MAY have a short life span
- XML Elements have attributes
- XML documents resemble traditional relational and object database data.

XML

- Preserves abstract data
- Prevents data from getting mixed up with presentation layer
- Proposed data interchange format for web data

KEY POINT - XML treats web documents as data, not just format

XML is license-free, platform-independent, and well supported

XML Tagging can be Used For...

- Structured data (e.g. databases)
- Semi-structured data (e.g. text)
- Unstructured data (e.g. images, multimedia)

… in both web and non-web scenarios

Or Another Way to Think of it ...

- XML for Content Management
- XML for EDI and e-Commerce
- XML for Enterprise Application Integration

The XML difference

documents are data

...

which can be presented as documents

The XML Difference ...

XML Delivers Data to the Client
while
HTML Delivers Presentation Only

Example - e-Invoice

e-business

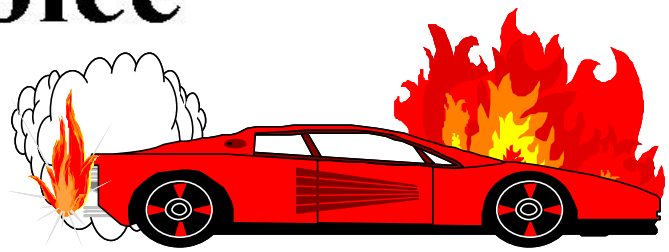
E-Invoice - Netscape

File Edit View Go Communicator Help

Back Forward Reload Home Search Guide Print Security Stop

Bookmarks Location: file:///c|/data/presentations/inv-html.htm

Automobile E-Invoice



Purchaser:

Eric Severson, Executive Consultant
IBM Corporation

Joe's Hot Deals

Item Purchased:

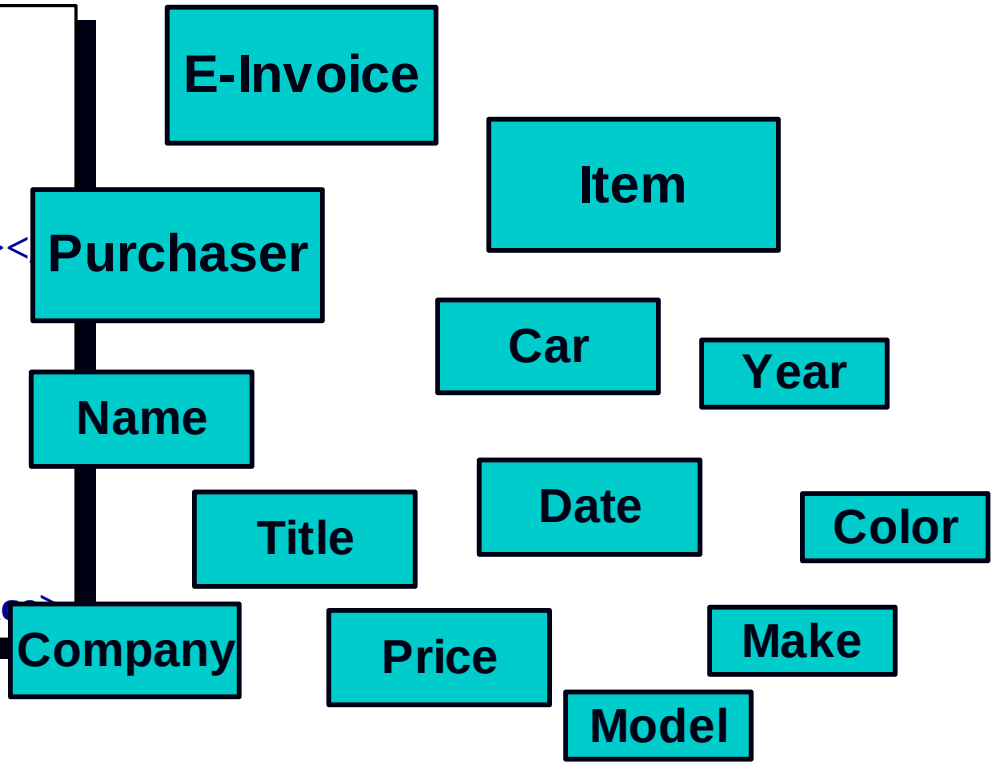
1997 Black Jeep Wrangler
Purchase Date: 7/1/96
Purchase Price: $20,000

Document: Done

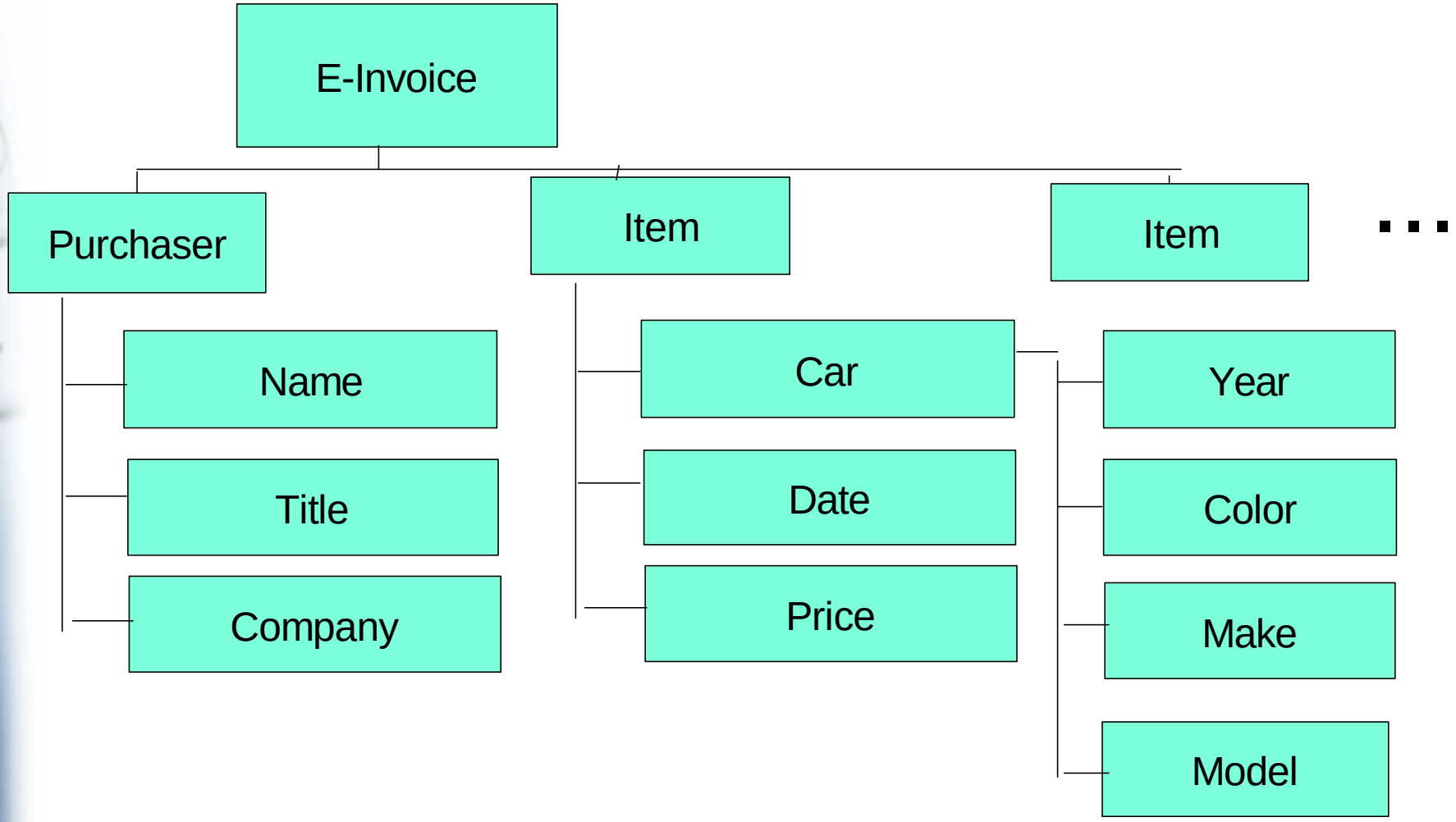
Start | Microsoft PowerPoint - [... | inv-html.htm - Notepad | E-Invoice : file:///C|/Da... | E-Invoice - Netscape | 7:39 AM

XML is made up of Information Objects

```
<E-Invoice  Dealer="Joe's Hot Deals">
<Purchaser>
<Name>Eric Severson</Name>
<Title>Executive Consultant</Title>
<Company>IBM Corporation</Company><
<Item><Car>
<Year>1997</Year>
<Color>Black</Color>
<Make>Jeep</Make>
<Model>Wrangler</Model></Car>
<Date>7/1/96</Date>
<Price>$20,000</Price></Item></E-Invoic
```

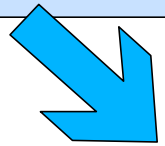


..which are defined in a structure with DTDs



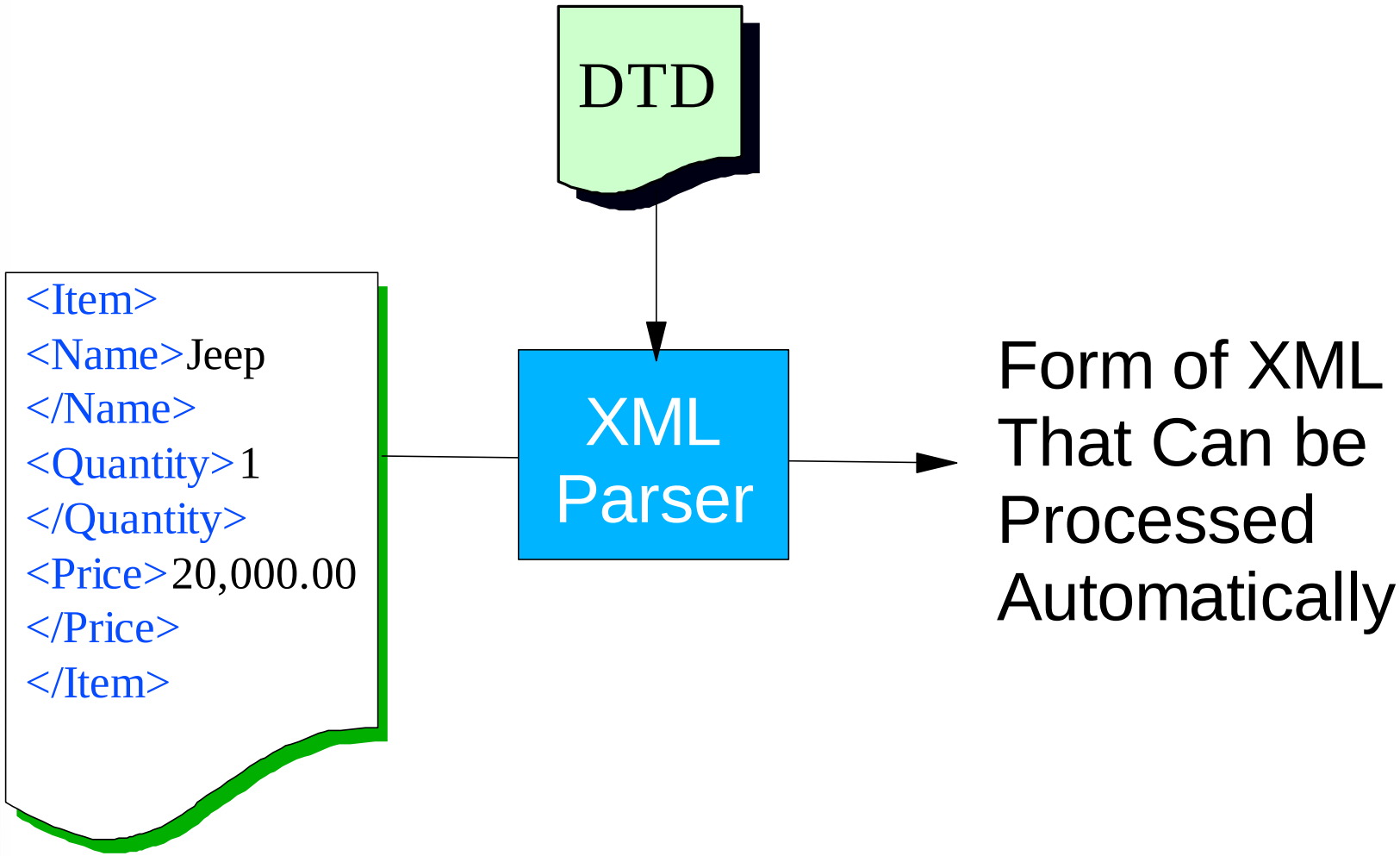
Relational Data in XML Form

Name	Quantity	Price
Jeep	1	$40,000

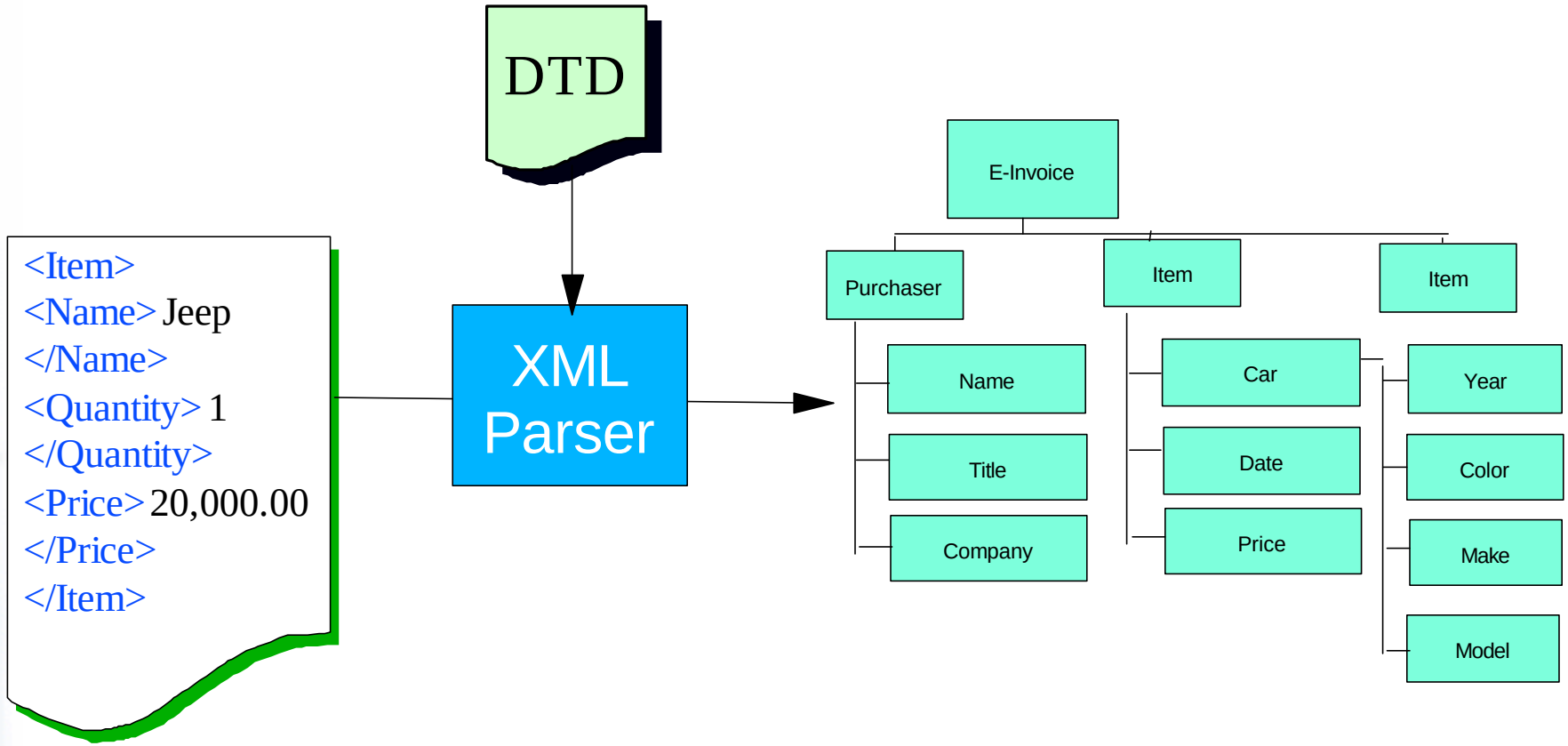


```
<Item>
<Name>Jeep</Name>
<Quantity>1</Quantity>
<Price>20,000.00</Price>
</Item>
```

XML Data is processed with "Parsers"

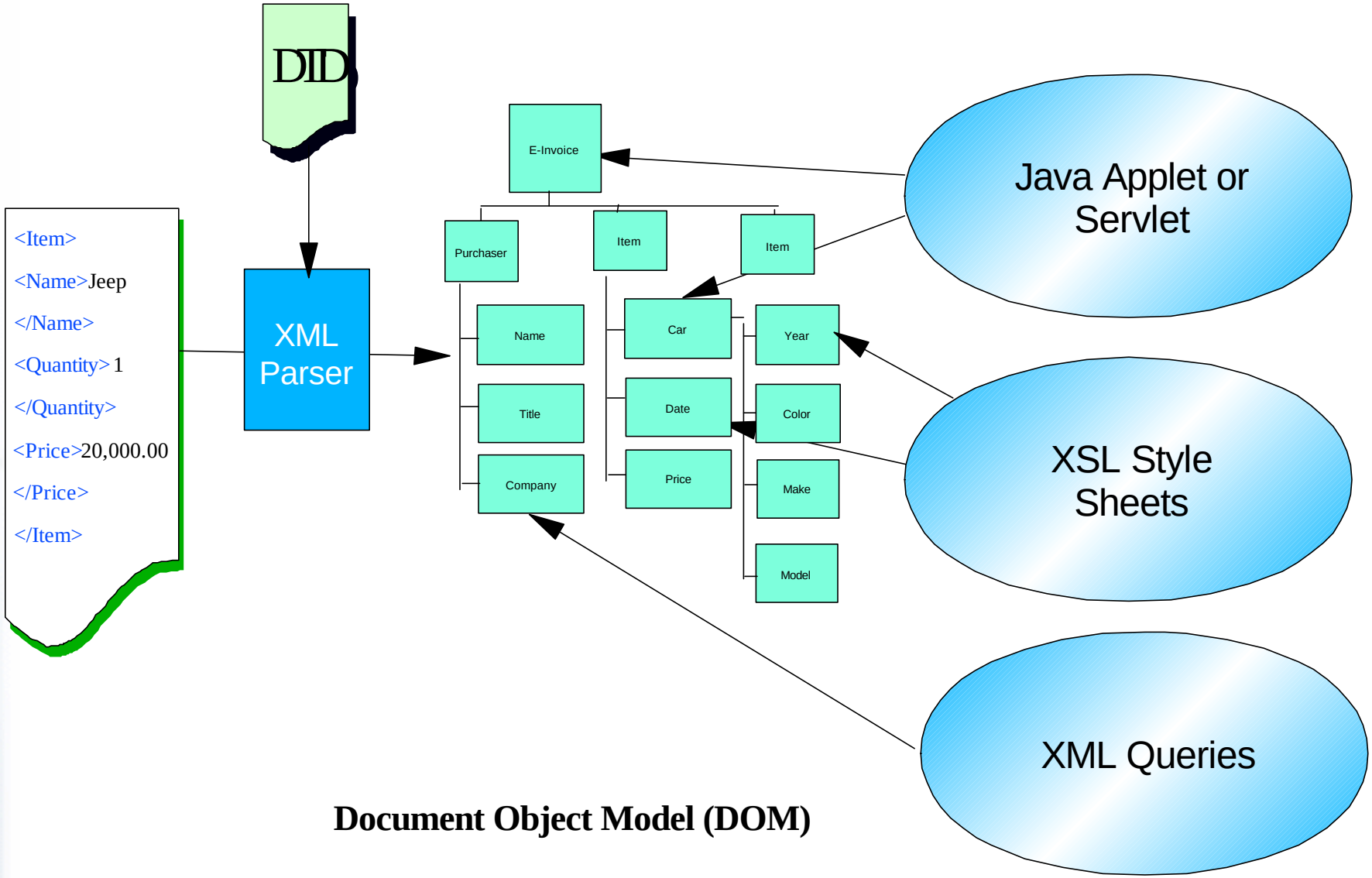


Which produce an API into the XML Structure



Document Object Model (DOM)

Accessible to Java, Stylesheets, and Queries



Document Object Model (DOM)

Java and XML Work Together

- **Java is the means to produce portable code**
- **XML is the means to produce portable data**

What others are saying ...

"XML is the single most important thing to happen to e-commerce yet, and possibly the more important than the internet"

Randall Whiting

CommerceNet CEO

InfoWorld Electric 11/17/98

“It is enterprise data integration applications, rather than mere Web publishing, that will prove the greatest strength of XML.”

Analyst JP Morgenthal

NC.Focus

Internet Week 2/8/99

Challenges in Using XML

- Need for Common XML Vocabularies (Data Dictionaries)
- Techno-hype has generated confusion
- Too many standards organizations all trying to do the same thing
- Tools coming fast, but not mature yet

However, products are emerging with XML Support in them (more on this later...)

XML Tools being created

- Tools to create XML documents from relational data -- with minimal programming
- Tools to create relational data from XML documents -- with minimal programming
- Tools ("parsers"), available on the server and the client, that read and validate XML documents
- Tools to create HTML pages from parsed XML; on the server and the client
- Tools, on the client, to create XML documents from data entered on HTML pages

Standards Groups Involving XML

- W3C
- XML.ORG
- BizTalk
- SAX
- OMG
- OASIS
- IETF
- DISA
- HL7
- ANSI
- OBI
- ACORD
- ECML
- VICS
- FpML
- XFRML
- RosettaNet
- Dublin Core
- VXML Forum
- CommerceNet
- OpenApplications
- ebXML
- ActiveStore

Who is working on XML Today ?

- IBM –
- Microsoft and Netscape
- Sun and Java
- Oracle and other database vendors
- CommerceNet and other EDI stakeholders
- Web content management vendors
- SAP and other ERP vendors
- All the former SGML vendors
- A host of others ....

IBM & Apache Support XML Standards

- IBM contributes key XML technologies (9-Nov-99)
  - IBM XML Parsers for Java & C++ and LotusXSL processor form code base
  - Formation of Apache open source project - http://xml.apache.org/
- Open source accelerates vendor-neutral XML standards
  - Openness of the Internet supported by public standards
  - Non-proprietary implementation of W3C recommendations
  - Public participation through code contribution
  - Strong leadership: Apache, IBM, Lotus, DataChannel, ExOffice, Sun, and Textuality
- Ongoing demonstration of IBM's dedicated support of industry standards
  - W3C, OASIS, XML.org
  - XML, Java, Linux
- www.apache.org

e-business

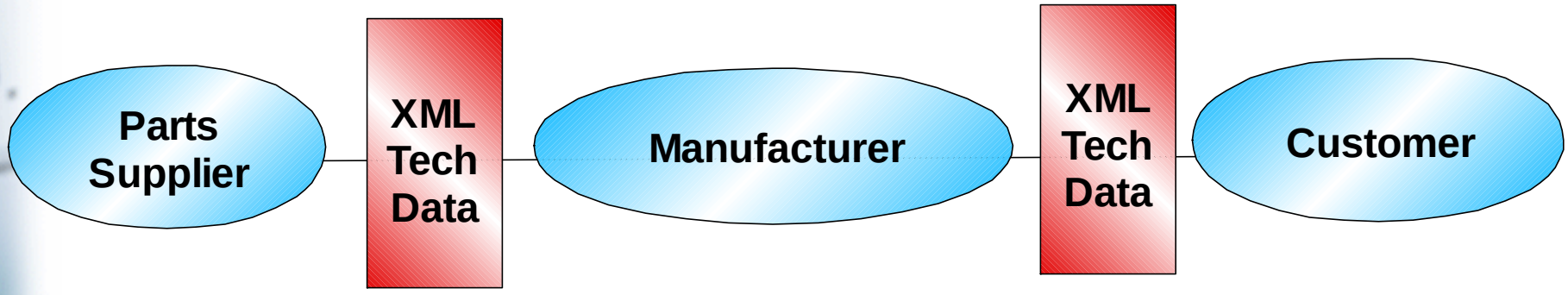
Who will be using it ?

- Airlines
- Automotive
- Financial Services
- Insurance
- High-Tech Manufacturing
- Banking
- Healthcare
- Telecom
- Information Publishing
- AND RETAIL !

e-business

How will it be used ?

Integration Across Enterprises



Application to Application Interfraces

Consider this ...

- CRM -- enterprise manages a centralized master db
- SCM -- RosettaNet.Org is a consortium creating an XML based trading process for IT supply chain
- ERP -- SAP moving to XML interface rather than native

Standardizing the XML tag definitions (DTDs) for these application domains will allow standard retail store application connections

XML and Retail Standards Organizations

- UCC effort
- ActiveStore Retail Business Interfaces
- ARTS REDX effort
- VICS
- and more sure to come

Why should you care ?

- Software Developers
  - Need to get up to speed on tools and architecture
  - Need to monitor standards groups to get standards that are workable for your applications
- Distributors
  - Your Customers may want to change their interface to you -- this whole eBusiness stuff
- Sales Teams
  - New Technology that is catching your customers eyes
  - What can you sell ?

What does XML mean to retailers ?

IBM products that exploit XML technology

- DB2 Universal Database 6.1
  - uses XML for web publishing and content management
  - data-exchange format
- IBM SecureWay Host On-Demand v4
  - XML for macro scripting
- IBM WebSphere Application Server
  - IBM XML4J - XML parser written for Java
- Catalog Architect
  - XML support allows ISVs to combine their products with Catalog Architect
- MQSeries
  - supports XML and Java

IBM Resources to get you started

- developerWorks
- jStart program
  - For more info contact jStart@us.ibm.com
- alphaWorks
  - XML parsers in Java and C++
  - Lotus XSL processor
  - www.ibm.com/alphaworks
- XML Zone on the developerWorks web site
  - www.ibm.com/developer/xml

Some Definitions

- XML - Extensible Markup Language
  - XML 1.0 - specification that defines what tags and attributes are
- XSL - formating rules, XML stylesheets
  - advanced language for expressing style sheets
- Xlink and Xpointer - ways to locate XML data
  - Xlink -still in development
    - -- describes a standard way to add hyperlinks to an XML file
  - XPointer & XFragments - also still in development - syntaxes for pointing to parts of an XML document

More Definitions

- CSS - style sheet language
- XML namespaces - a spec that describes how you can associate a URL with every single tag and attirbute in an XML document.
- DOM - Document Object Model
  - standard set of function calls for manipulating XML (and HTML) files from a programming language
- SAX - Simple API for XML
- XML Schemas - Data typing for XML
  - help developers to precisely define their own XML-based formats
- XQL - XML Query Language

Need More Information ?

- XML Complete, by Steven Holzner, McGraw Hill, ISBN 0-07-913702-4
- Building Corporate Portals with XML, by Clive Finkelstein & Peter Aiken, McGraw Hill, ISBN 0-07-913705-9
- Designing XML Internet Applications, by Michael Leventhal, David Lewis, Matthew Fuchs, Prentice Hall PTR, ISBN 0-13-616822-1
- The XML Handbook, by Charles F. Goldfarb & Paul Prescod, Prentice Hall PTR, IXBN 0-13-081152-1
- WWW.W3.ORG - W3C consortium
- www.xml.org