News   Vol.2(1) June 2000

An overview of XML

Margaret van Steenderen
Margaret.vansteenderen@commercequest.co.za

Why is XML seen as so important, and why is it fast becoming the industry accepted standard for data transfer? The first column looks at what XML is and why it was developed and the next looks at the impact of XML. Thereafter, some packages that use XML will be reviewed with a focus on the business (and especially information sharing) benefits.

XML stands for eXtensible Markup Language, but what does that really mean? XML is a standard. Just as HTML is the standard that made formatting and displaying pages on the Web viable, so XML is the standard to make data sharing and transferring viable.

As with HTML, XML is by no means Internet dependent. Although its main application seems to be transferring information on the Internet (because the Internet is being used for B2B transactions), it is perfect for data transfer between systems, regardless of the network being used.

top

Why use XML?

XML is a way of describing data in a platform and application independent manner. It allows for the description of complex data in such a way that the information is self-describing, practically removing any problems with change or conflicts over the format of information, especially significant when looking at server-to-server transfer of information.

Or, in one line:
XML provides self-defining data.

top

An example

Why can XML be defined as 'self-defining data'? Here is an example:

```
<customer>
  <name>Margaret</name>
  <surname>van Steenderen</surname>
  <address>
  <line_1>P O Box 1074</line_1>
  <city>Cresta</city>
  <pcode>2118</pcode>
  </address>
</customer>
```

The <enclosed words> are the tags. As you can see, a tag's beginning definition is in the brackets <tag>, and the tag ends with a slash in the brackets </tag>. The data are between the two tags. This makes it really easy to see what data belong to what tags. If the above was a flat file, it could look something like the following:

Margaret|van Steenderen|P O Box 1074|Cresta||2118

There is no real hint as to what the data fields are. Of course, if there are missing fields these would be represented by a form of punctuation or other special character, again giving no hint as to which they are. With XML a missing field may be represented by a tag that has no attribute value, for example <empty></empty> or the short-hand version <empty/>

HTML vs XML

According to Melnick (2000:Online), 'HTML limits the ability of the average Web page to share data or interact with other Web sites. This is largely because HTML has been called upon to do many things it was never designed to do. However, XML has come to the rescue, offering creativity and data exchange almost as limitless as HTML itself is limiting.'

HTML is the markup language for how documents are displayed. It controls how the document is formatted and what it looks like on Web browsers. It has nothing to do with data or information types or formats. HTML can almost be seen as the word processor of the web. With the continuing growth in e-commerce and business-to-business applications, there is a need for a standard language that defines and structures data across platforms and that is neither development language nor database dependent. XML fills this gap. It is important to note that XML is not a replacement for HTML. HTML is still 'an excellent tool for displaying hypermedia documents across a network. … XML is designed for electronic information providers who want to do things that HTML isn't designed for.' (Textuality, 2000:Online).

Who supports XML?

XML is headed by the World Wide Web Consortium (the W3C). It is an open standard and therefore public property. A few of the many supporters of XML are mentioned below:

- ISO, W3C
- Microsoft (Office2000, IE4/5, VisualStudio, VBScript…)
- Oracle (8i, iFS)
- Platinum repository
- IBM DB2 (Project Garlic), Lotus
- SAP R3
- Sun (Security data exchange using XML and JavaBeans)
- CommerceQuest (XMSeries)
- Software AG (Tamino)
- And more …

Wrapping up

The above is a very small selection of supporters of XML and a brief overview of XML. There are many tools available for XML, some free, some not (especially those that add significant value). XML has already had a big impact on business, and, like the Web, it will be a case of 'use it or lose the business'.

References

Melnick, J. 2000. Globalizing e-commerce. [Online]. Available WWW: http://www-4.ibm.com/software/developer/library/globalsoft.html

Textuality. 2000. Extensible Markup Language (XML). [Online]. Available WWW: http://www.textuality.com/xml/

Other Useful Web Sites

Learning XML and general information:

- http://developer.earthweb.com/journal/techworkshop/xmlport.html (Very good technical tutorials)
- http://www.ucc.ie/xml/ (FAQ)
- http://www.xml.org/ (Resources / News)
- http://www.xmlmag.com/ (A free online XML magazine)
- http://www.w3.org/XML/ (Provides the XML specification and a lot of other useful information)

Disclaimer