

XML - The e-Business Glue

Mike Bromwich, Technical Director of PDMS

June 2000

Welcome technophobes! In this article, I present XML- but please do not turn the page fearing a list of acronyms and a technical explanation of how to build a new generation web site.

XML stands for eXtensible Markup Language. The concept is so simple that it is difficult to appreciate why it has taken so long to arrive. XML introduces a fundamental set of practical business concepts. This is a case of a technology where the technicalities are very basic and even mundane. To grasp the XML philosophy, we can ignore the internal workings totally.

In a nutshell, XML is a way of describing the format of data. There are plenty of data formats in common use; many are horizontally applicable to several diverse areas, such as a word processor file. Other formats have a vertical bias; for example GEDCOM is used to store genealogy information. These formats are very different. Trying to load a GEDCOM file into your word processor will result in a collection of characters that appear to contain no useful information at all. XML takes a step back from problems such as this, and specifies a mechanism for DIY data formats, safe in the knowledge that it will result in interoperability.

To make e-business a reality, you need to be able to speak to your customers, suppliers, partners and employees electronically, efficiently, and with minimal human intervention. More importantly, their computer systems need to be able to understand what you are saying, and respond intelligently. Although each party will make their own choices of computer system and software, they must be able to integrate into the resulting e-business framework. In the XML world, the key to this interoperability is a schema.

The easiest way to understand the role of a schema is to take an example. When your accounts department receives an invoice from a new supplier, it can be interpreted and processed without any instruction from the sender. This is because they understand what an invoice is; it has a tax point date, an invoice number, lines containing quantity, description etc., and one or more totals. The differences in layout of an invoice from one supplier to the next are not important. However, try scanning these invoices directly into your accounts software, and the results would be unreliable at best. Even if suppliers send an invoice electronically by e-mail, you could not automatically process it unless your software knew exactly what piece of information to expect where.

XML addresses this problem by separating the content of the invoice from the presentation. The content of the invoice is the useful information, including data such as the tax point date and invoice number. These pieces of information will be labelled according to an XML schema, a specification previously agreed between the parties. There are standard schemas being developed for all common business documents such as invoices, statements, and purchase

orders. The second body of information, the presentation, describes how to present the content so it appears how the sender intended. This may include layout, font and logo information in accordance with your company's corporate image. Typically, this presentation information will be supplied as another XML document, known as a stylesheet.

In processing the invoice automatically and electronically, your accounts package can safely ignore the presentation information, and can usefully digest the content. If the content includes any additional information which is not required, this can also be automatically be recognised and skipped. XML is self-describing, meaning that software can identify the various important information fields and discount any additional information that the software does not recognise or require.

As a result of using an XML document to represent our invoice, it is now possible to dispatch, transport, receive and process documents electronically, providing vast improvements in efficiency, confidentiality and accuracy.

Although I have used an invoice as an example of an XML application, the true power which XML promises is the ability to replace almost any paper document transaction with an electronic equivalent. It will handle the exchange automatically through all stages of the e-business chain. The XML systems transporting our 'paperwork' may also be processing CML (Chemical Markup Language) files or any other XML-based document. An XML schema can be defined for your industry today, and it can be processed by any existing XML system tomorrow. As long as you define which schema will be used, you can exchange documents electronically with any other party with very little effort.

The key that will enable us to make the most of XML is the adoption of compatible schemas, and there are several initiatives in place to encourage this. One of the highest profile ventures is BizTalk, started by Microsoft. It houses hundreds of schemas covering everything from exchanging business cards to chemical formulations. It has gained the acceptance of many of the influential players in industry, such as SAP and CommerceOne, Boeing and BP/Amoco. This initiative provided an early clue to the strategic importance Microsoft places on XML; indeed it is a cornerstone of the new Windows 2000 operating system.

The introduction of XML schemas provides a springboard for the adoption of business-to-business and business-to-consumer electronic document exchange. The schemas support encryption and digital signatures and there is a gradual world-wide introduction of enabling legislation, such as the Isle of Man Electronic Transactions Bill. Current estimates are that over 80% of paper documents exchanged between businesses are produced by computer, transported by post, only to be entered into another computer. Some of the newest technology available is currently constrained by one of the oldest technologies in use, and so XML is almost certain to have a huge influence if it lives up to its promises.

Published in Money Media, June 2000