

Effective strategies for integrating businesses on the Net using XML

David Burdett

Commerce One, Pleasanton, USA

david.burdett@commerceone.com

<http://www.commerceone.com>

Abstract:

Many businesses plan to use eCommerce to connect their accounting and ERP systems to their business partners. However connecting systems together is neither easy nor straightforward. Unless systems can communicate properly and effectively, the anticipated business benefits will not arise. Many people agree that XML has a major role to play in making this work. But in practice what must you do make it work?

Commerce One has to address this very real problem with the many companies whose systems we interconnect using Commerce One's MarketSite business-to-business portal.

The first part of this talk will explain the problems that need to be solved covering: Documents, Choreography, Messaging and Transport.

The second part of the talk will address, in more detail, some of the standards initiatives that are addressing the barriers described in the first part such as Microsoft's BizTalk and SOAP, consortia based approaches such as ebXML and RosettaNet as well as standards groups such as the IETF and W3C.

In summary, from the first part of this presentation, managers who are responsible for or are planning to integrate their company's systems into an eCommerce environment will understand an approach they can follow to tackle the work and what are the major problems that they will need to overcome.

From the second part of the presentation managers understand the major initiatives that are addressing the integration problem and will be better positioned to identify which solution or approach will be best for them.

Introduction

This talk will identify and describe the key problems and issues that need to be addressed in order for businesses to do eCommerce with other business over the net using XML. Firstly it will describe the barriers faced in implementing interoperable eCommerce systems covering the four key requirements that you must get right if you want eCommerce systems to interoperate effectively: documents, choreography, messaging and transport and then illustrate the practical problems that need to be addressed from Commerce One experience. The second part of the talk will expand on "messaging". This will describe the requirements for delivering XML messages reliably over the internet and then compare and contrast developments under way in BizTalk, Soap, ebXML, RosettaNet and the IETF that are addressing the messaging problem.

Barriers to eCommerce

The whole rationale for eCommerce is that it lowers the operational costs for a business and improves the selling and buying opportunities. As a result profits should increase.

Lowering costs require:

more efficient supply chains through businesses being able to integrate their systems, for example ERP (Enterprise Resource Planning) systems, together so that they can interact in real time

businesses participating in communities, or marketplaces of similar business, to identify new partners they can trade with.

Firstly, let's look at some of the barriers to eCommerce that can prevent businesses from integrating their systems.

Connecting everyone together

Typically B2C (Business to Consumer) eCommerce (B2C) uses technology illustrated in the figure below.

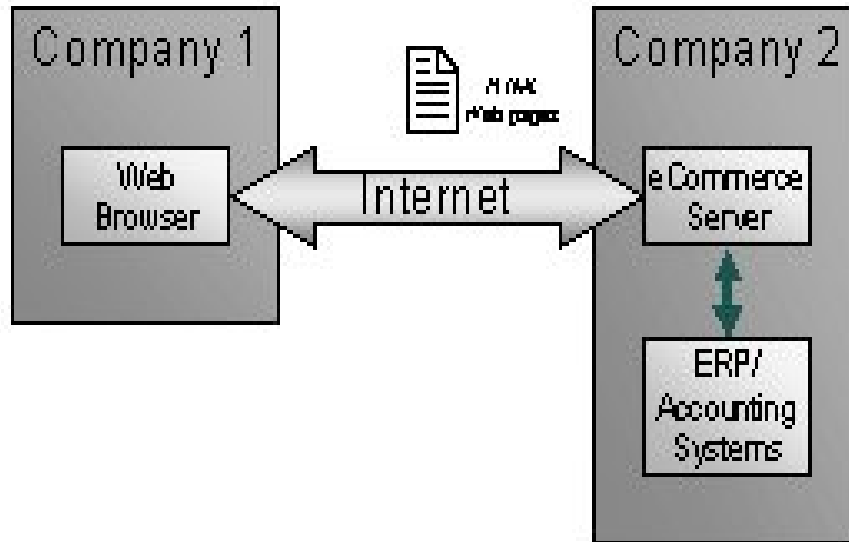


Figure 1. The Web Browser approach

The benefit of this approach for B2B eCommerce is that Company 1, using a web browser, does not need any special technology to connect with its partners. Company 2 can easily present web pages that can be used by company 1 to order goods or services. The harder job is for Company 2. It has to integrate their eCommerce or Web Server with their "back end" ERP or accounting systems. However some, particularly the smaller, companies are reluctant to integrate their Internet and accounting systems. As a result, data captured by Company 2 may, in practice, be re-entered manually into their ERP systems as well as Company 1.

A better approach is illustrated by the figure below.

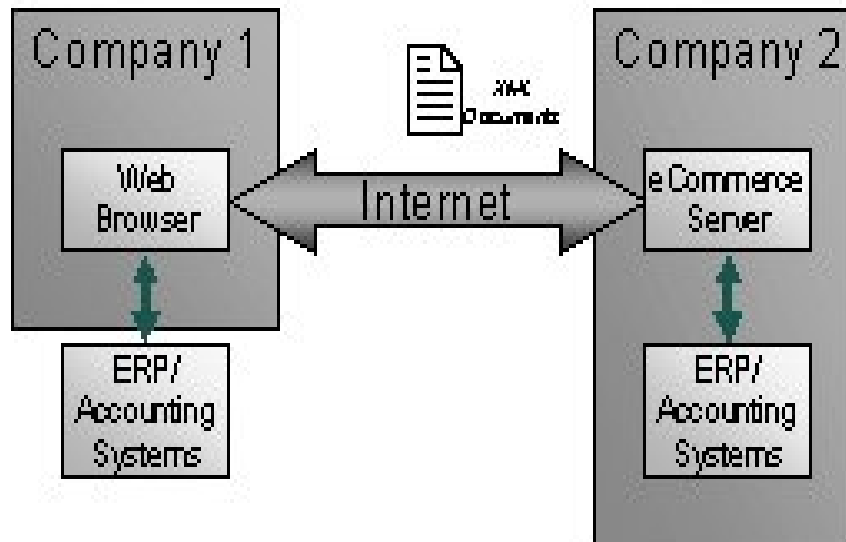


Figure 2. Integrating companies using XML

In this example both Company 1 and Company 2 have implemented XML technology to make their solutions integrate. This can offer significant cost benefits. Studies by Commerce One and others have shown that the cost of processing for example, Purchase Orders, completely in this way, results in cost reductions from over \$100 US to around \$8 US.

However exchanging documents is not completely straight forward since you need to know which format to use. Many ERP vendors are each defining their own versions of equivalent XML documents. For example consider the following two examples of XML.

```
<PurchaseOrder OrderNo="1234">
  <Buyer BuyerNo="AB24567"
    <Name="Smith & Co2/"
    <Address1>123 High St</Address1>
    <Address2>Epping Forest</Address2>
    <Zip>E15 2HQ</Zip>
  </Buyer>
  <Supplier>
  ...
  </Supplier>
  <OrderItem>
  ...
  </OrderItem>
  <Tax>
    <TaxType>VAT</TaxType>
    <TaxPercent>17.5</TaxPercent>
    ...
  </Tax>
  ...
</PurchaseOrder>

<PurchaseOrder>
  <Party Type="Buyer">
    <Reference>AB24567</Reference>
    <Name>Smith & Co</Name>
    <Street>123 High St</Street>
    <Town>Epping Forest</Town>
    <PostCode>E15 2HQ</PostCode>
  </Party>
  <Party Type="Supplier">
  ...
  </Party>
  <OrdNo>1234<OrdNo>
  <OrderItem>
  ...
  </OrderItem>
  <Tax Type="VAT" Percent="17.5"... />
  ...
</Purchase Order>
```

Figure 3. Two equivalent XML documents

Purchase Order Sample 1
Purchase Order Sample 2

Even though these two sample purchase orders contain **identical** information they have different structures. This means that if one company is using one format and a second company the other, then one or other of the companies will need to map their XML Purchase Order documents to the alternative format.

The conclusion to be drawn from this is that XML, on its own, will not provide a complete answer to the problem of defining documents for eCommerce.

Mapping document formats

This leads to the first barrier that must be overcome before eCommerce can be successful: i.e. it must be possible to map one document format to another. Consider the figure below where you want to map an EDI transaction to an equivalent one in XML.

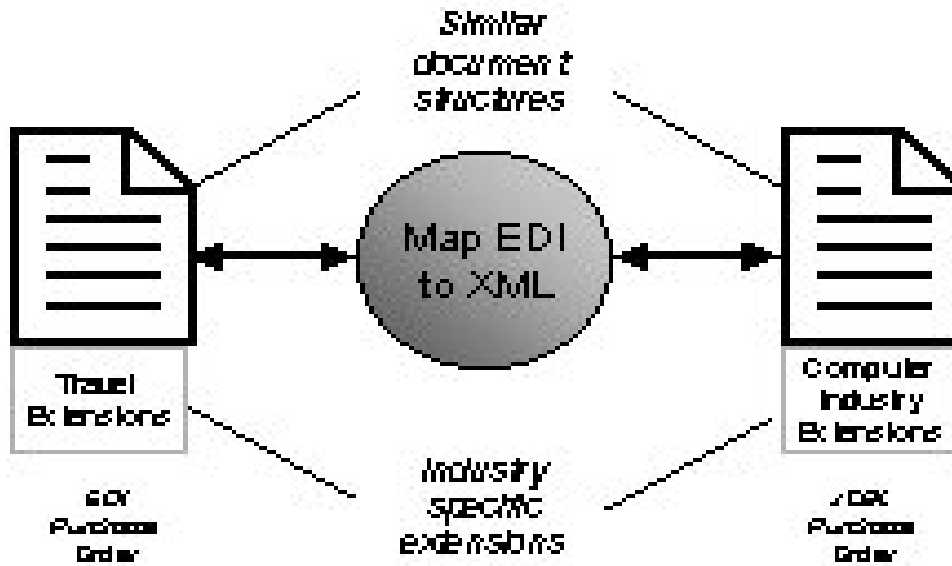


Figure 4. Mapping documents
Mapping an EDI document to an XML document

Firstly, for document mapping to work, the two documents that are being mapped must have essentially similar document structures and contains the same or similar data. For example in Commerce One, we designed xCBL after a careful analysis of the equivalent EDI documents. As a result it is a relatively straightforward process to map from an EDI document to its equivalent in xCBL.

However there is also a need to provide support for particular industries, for example, the additional data required on a Purchase Order to describe the characteristics of a memory chip ordered from a semi-conductor manufacturer are quite different from the data required to describe an airline flight ordered from a travel agent.

Historically, EDI has addressed the needs of different industries by gradually adding more and more to the base EDI transaction definition until no-one can directly use the EDI transaction in its entirety. This means that, companies that want to exchange EDI documents have to agree the subset of the complete EDI transaction data that they will use. As a result it is very rare that two EDI implementations are the same and each needs to be tailored, at significant cost, to the needs of each situation. This presents a barrier to the adoption of EDI systems particularly for smaller companies.

XML, and particularly XML developed using schema languages such as Commerce One's SOX, or the W3C's XSDL, take a different approach where the common subset of data for a document is identified to create a "base" XML document definition. This is then "extended" with additional data elements to meet the

need of particular industries. Effectively it allows a group of travel of companies or their industry association to say "my purchase order is like the standard purchase order except that it has these extra fields."

Matching document choreographies

The second barrier that you must be removed is associated with document choreography. This is the area that is the most hard to handle as it requires that the business processes in companies that want to carry out eCommerce are aligned so that business documents are exchanged between them in a sequence that both can understand.

For example consider the document choreography illustrated by Figure 5 below.

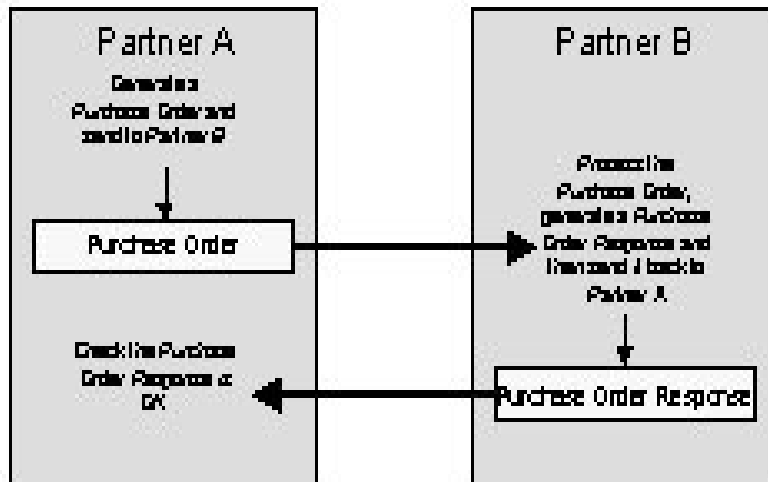


Figure 5. Simple document choreography

In this very simple choreography, Partner B responds to Purchase Orders it receives by sending back a Purchase Order Response that, typically, describes what is in stock what is not.

On the other hand if Partner B's systems could only respond with an invoice instead, then Partner A's systems would not be able to handle it as they would not be expecting it, and so the two Partners cannot carry out eCommerce.

This is a very simple example and in practice, document choreographies can be far more complex. It is therefore important to discover and understand what choreographies two businesses use to communicate one with another and make sure that the business processes at each hand can support it.

One of the major contributions of organizations such as RosettaNet is to develop standard document choreographies and document definitions that make it much easier to develop inter-operable solutions.

Messaging compatibility

The next barrier that must be overcome is in the area of Messaging. At its most simple level, messaging is about sending a document from one place to another over the Internet. However this only works if the outer

electronic envelope that wraps a business document is understood by both the sender and the receiver of the message.

To take an analogy from the real world, consider the figure below that illustrates different ways of sending paper or other physical documents between two places.

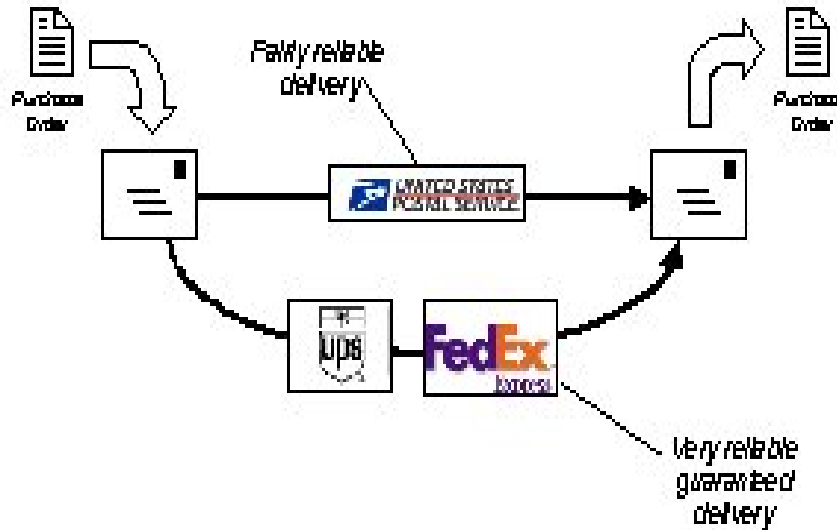


Figure 6. Conventional postal services
Real world message delivery

The most popular way of sending documents to someone is by using national postal services such as the USPS (United States Postal Services) . However if you want to send a letter quickly and reliably from one place to another, then you make use of companies such as FedEx or UPS who provide reliable courier services. The USPS, FedEx and UPS, though, have their own forms that **must** be used to address their packages whether it is the particular style in which you write the destination address on an ordinary envelope or the multi-part pre-printed forms used by FedEx and UPS to enter the same information. If you don't use the correct "form" for the postal service that is transporting them, they can't be properly handled.

The same applies in the electronic world as illustrated by the figure below.

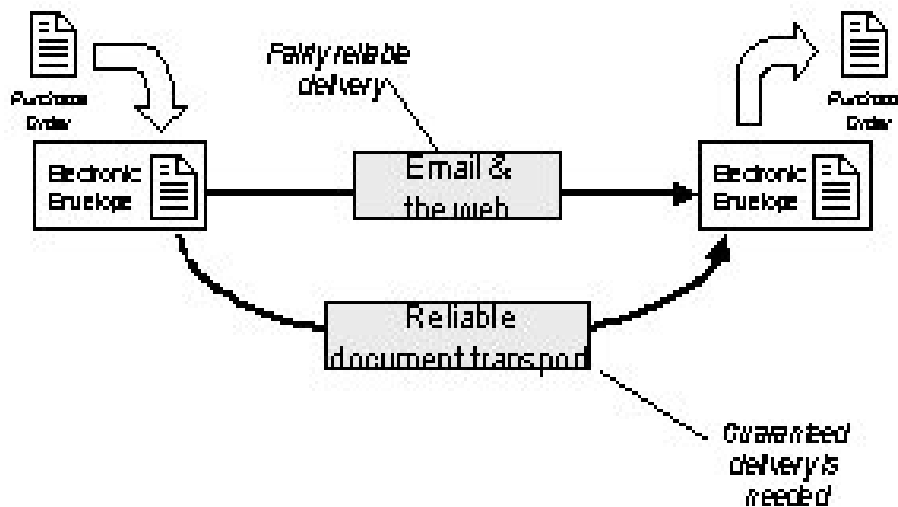


Figure 7. Internet "postal systems"
Internet message delivery

Just as in the, real world, the document you want to send, for example a Purchase Order defined using XML, must be placed inside an electronic envelope before sending on its way. Also, just as in the real world, the structure of the electronic envelope needs to be understood at each end and everywhere in-between. Currently there is no universal standard for electronic envelopes. This means that frequently you must use proprietary standards or standards with limited adoption to connect.

There is another problem in that internet transport protocols such as SMTP for email or HTTP for sending web pages, are not completely reliable. There is no universal electronic equivalent on the net of FedEx or UPS. For B2B eCommerce this is simply not good enough. Sometimes "just maybe" delivery of critical business documents is unacceptable.

Commerce One's MarketSite however has solved this problem by implementing a reliable message delivery mechanism inside Commerce One's XCC. This means that, when you need to use it, you can be sure that your documents will get through, and, if delivery is just not possible - for example the destination server is down - then you will be informed that the delivery failed.

However Commerce One recognizes that not everyone will want to use XCC and therefore we support other methods of connection that provide varying degrees of reliability. This is illustrated by the figure below.

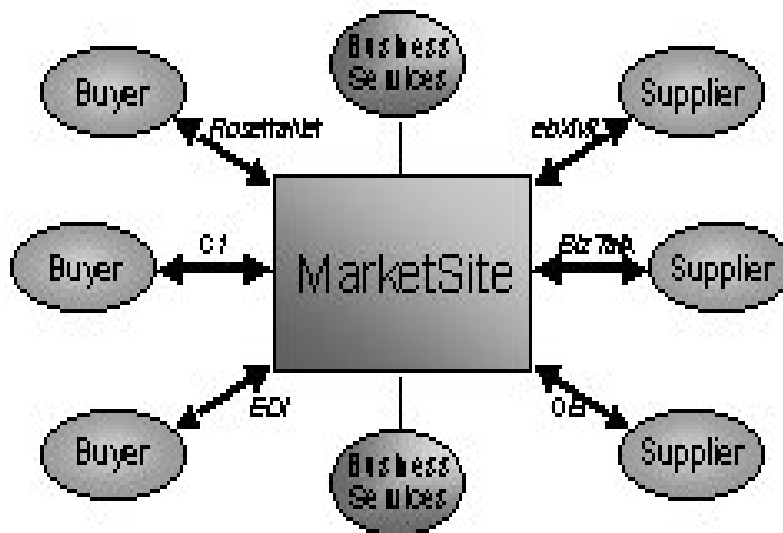


Figure 8. Commerce One - the open portal

Commerce One is committed to the implementation and support of open standards that provide our partners with a choice on how to connect to our products and are active participants in several standards initiatives such as RosettaNet, ebXML and Microsoft's BizTalk. These initiatives are discussed in more detail later in this paper.

Transport protocol compatibility

The barrier that needs to be overcome is in the area of Transport Protocol compatibility. In this case the underlying transport protocol that is used such as HTTP or SMTP must be the same at each end. One way

of thinking of transport protocols is to think of them as the electronic equivalent of the trucks that are used to transport documents physically from one place to another. In the real world compatibility is not usually a problem since trucks and roads have been around for a long time and all the incompatibility problems have been ironed out - e.g. everyone in one country drives on the same side of the road!

Fortunately the same is true in the virtual world. HTTP and SMTP have been around for a long time therefore compatibility problems have been resolved.

In the real world there is a choice of transport methods that you can use, for example train, truck, air, or ship - or some combination of them all that is used.

The same situation exists in the virtual world in that some companies, especially smaller ones, might only be able to support eCommerce by sending their documents using email. The big companies, on the other hand are more likely to want to use HTTP since it is faster and more immediate. This can result in a "mismatch" between transport protocols.

Commerce One helps in this area by providing support for different types of transport protocol. For example one company could send a document to Commerce One's Market Site using HTTP and then Market Site would transport it to its destination using SMTP or vice versa.

This concludes the discussion on connecting everyone together.

Participating in communities

Helping companies find each other

Even when companies have overcome the barriers that allow them to be connected with one another, there is still a need for companies to discover other businesses they can trade with on the net.

You probably know that somewhere out there is a buyer who wants to buy your goods, as well as a supplier who can provide you with the competitive products and services that you need to be successful.

Typically, if you want to find someone on the web, you use search engines such as those provided by Yahoo! and others. But these are very hit and miss in that you have to put in vague searches that probably bring back far too many results for you to be able to identify what you really need..

If companies participate in communities, then much more structured queries are possible where:

- a supplier registers their products and services with the community, then

- buyers perform structured queries using XML that provides much richer and more accurate results to the query.

The same can happen in reverse where the buyer registers their need for a service and suppliers can then respond. This can then easily be taken to the next step where a buyer can carry out reverse auctions or RFQs (Request for Quotes) using the community and suppliers can then make their offers in a much faster way and without the effort of having to go physically to the buyer's premises.

Full blown auctions are also easy to do where excess inventory can be sold-off on-line.

Internet standards initiatives

There are many different standards initiatives that, one way or another, focus on eCommerce and address the barriers described earlier in this paper, such as:

Microsoft's BizTalk (<http://www.biztalk.org>)

ebXML - electronic business XML (<http://www.ebxml.org>)

RosettaNet - (<http://www.rosettanet.org>)

eCo - (<http://www.commercenet.org/eco>)

IETF - (<http://www.ietf.org>)

W3C - (<http://www.w3.org>)

To provide some clarity on these initiatives, this part of the paper reviews them so that the reader can better understand what they are trying to achieve and how they might help in implementation of eCommerce solutions. There are in fact many more initiatives than these that space does not allow be covered here. It is also a very rapidly moving field so the information provided here will probably soon be out of date. However the information provided is believed to be correct at the time of writing - late March 2000.

Commerce One is involved in all the above initiatives since we know that adoption of standards is important to our customers who want to be able to carry out eCommerce using practical but effective open standards.

All of the above initiatives address one or more of the barriers that have been described earlier by providing technology that can be grouped under one of four headings:

Common XML Document Definitions - that attempt to develop "standard" XML definitions for the data from which business documents can be created

Messaging Standards - that tackle the problem of inter-operable exchange of documents over the Internet

Architecture/Discovery Standards - that allow you to find out information about businesses and the eCommerce services they provide

Registries and Repositories - that allow you to discover information about XML document definitions and other related information

Each of these is discussed in more detail below. Please note these comments are the opinion of the author and no-one else!

Common XML Document Definitions

Common XML Document Definitions addresses the development of "Standard" XML definitions for business data. For example you could define an XML structure for "Name and Address", "Payment Amount", "Tax", "Line Item", etc and then assemble them together to create recognizable business documents such as a "Purchase Order", "Invoice", "Advanced Ship Notice", etc.

The benefit of taking a "modular" approach is that implementers can develop a standard piece of code to manipulate, validate or display, for example, a Name and Address, which can then be reused in ANY software that processes ANY document that uses the "standard" structure.

If you also take a minimalist approach to the design of your documents you can, using Schema languages (see description earlier in this paper), extend the definition of the individual components within the document to meet the need of particular vertical industries. For example you can extend the standard "Line Item" for the travel industry to include, flight information such as flight number, start/end airports, departure/arrival times, ticket no, etc.

In many ways extending documents in this way is very similar to Object Oriented programming "class inheritance" mechanisms. It means that you can quite easily write one piece of code that can handle the base level document as well as any document that is based on the base document but has been extended with extra information. The code can then carry the extended information, but safely ignore it until it reaches some software that recognizes the extensions and knows how to handle them.

Initiatives

There are several initiatives that are trying to develop common definitions. They can be split into two types:

- horizontal, cross industry initiatives, and
- vertical initiatives, that address the needs of a particular industry.

There are really only two horizontal initiatives: ebXML and xCBL. ebXML is a joint effort of the UN/CEFACT and OASIS.

ebXML - what is it?

The EDI community recognized that "XML is coming" and that it promises significant benefits in terms of better, easier and wider connectivity, particularly to the small- to medium-size enterprise.

Right now, though, EDI documents are at the opposite end of the spectrum from XML in terms of flexibility and understandability. You can't just do a "direct" translation from EDI to XML - it would be as bad as doing a "word-for-word" translation from German into English (or vice versa), because the grammar is different. You won't be able to easily understand the result.

Solving this problem is what ebXML is trying to do.

Essentially it is a jointly sponsored effort by OASIS, a non-profit XML vendor consortium, and UN/CEFACT, who "own" the EDIFACT standard. EDIFACT is the international version of the EDI standard; X12 is similar, but is US developed and owned.

What they hope to achieve is a combination of:

- the vast knowledge and expertise of the EDI community about **what data** needs to be exchanged between business to support e-commerce, and
- the XML expertise of OASIS members in developing "good" XML documents to produce a solution that will provide real business benefit.

ebXML - Core Components working group

The ebXML "Core Components" working group is developing standard elements. It's mission is to:

- develop guidelines/methodology for consistently building/deriving of core components,
- identify and develop reusable core components,
- define metadata for core business information model,
- define rules for extensibility,
- recommend procedures for approving core semantic elements,
- define algorithm/conventions for producing tag names,
- develop guidelines for bridging core elements from EDI.

More from http://www.ebxml.org/project_teams/core_components.htm (note underscores between spaces).

Commerce One is developing the reference implementation for Core Components.

xCBL - XML Common Business Library

xCBL is now in its second version and is the first open XML specification for the cross-industry exchange of business documents such as product descriptions, purchase orders, invoices, and shipping schedules. xCBL 2.0 is a set of XML building blocks and a document framework that allows the creation of robust, reusable, XML documents for electronic commerce.

xCBL was based on a careful analysis of the equivalent EDI documents, this means that it is easy to map from EDI documents to their xCBL equivalents and is being used as one of the major inputs into the ebXML Core Components working group. xCBL is available on an unrestricted basis and can be downloaded from <http://www.commerceone.com/xml/>

xCBL is also available in all the existing schema formats: Commerce One's SOX and Microsoft's XDR. It will also be made available in the W3C's XSDL when that is available.

Vertical initiatives

There are several vertical initiatives that are addressing the needs of specific industries. For example

RosettaNet - <http://www.rosettanel.org> - for the Computer Manufacturing and Electronic Components industry

Open Travel Alliance - <http://www.openatravel.com> - for the travel industry

Financial Products Markup Language - <http://www.fpml.com/> - for the financial industry

What many of these initiatives have in common is that they define their own documents - and none of them are really the same.

There are many, many more. For example see <http://www.oasis-open.org/cover/xml.html#applications> which is a useful web site maintained by Robin Cover that lists over 200 different vertical initiatives of one kind or another, including such obscure ones as "Australia and New Zealand Land Information Council" and the "Mind Reading Markup Language"!

Messaging standards

There are three main messaging standards initiatives:

Microsoft's BizTalk Framework - <http://www.biztalk.org>

The ebXML Transport working group - http://www.ebxml.org/project_teams/transport.htm

RosettaNet's Implementation Framework (RNIF) - <http://www.rosettanel.org>

BizTalk Framework

The BizTalk Framework is Microsoft's method of wrapping XML documents with a header that identifies where the document should be sent and is designed to provide the structure for the development of interoperability in a platform and technology neutral manner.

It provides specifications for the design and development of XML-based solutions for communication between applications and organizations by leveraging standard internet technologies such as XML, XML-Schema, and MIME.

A key requirement driving the BizTalk Framework specifications is the need to enable applications to securely and reliably exchange business information in XML. This may be implemented in standard envelope structures facilitating use of multiple transports, and supporting both XML and non-XML based business information.

The BizTalk Framework is open in that anyone can implement it. Microsoft plan to implement it as part of the BizTalk Server which is due out in the last half of 2000.

Commerce One is a member of the BizTalk Steering Committee.

ebXML Transport working group

The ebXML Transport working group is developing a protocol that has the objectives of

- enabling any party to carry out integrated eCommerce transactions with any other party anywhere in the world using their hardware and software vendor of choice

- persuading a wide variety of vendors to implement the approach

- not reinventing the wheel - re-use where possible

- enabling existing "messaging" solutions to "bridge" to the ebXML solution

- scaling from SMEs to large companies

- scaling from low power to high end solutions

Currently the ebXML Transport Working Group has published an "Overview and Requirements" document and is working on its first specification.

Commerce One is the editor of the ebXML Transport Working group.

RosettaNet

In order to solve the problem of how to transport XML documents over the net, the RosettaNet consortium developed a wrapper that defines how XML documents should be wrapped.

Commerce One is a member of the RosettaNet Implementation Framework development team.

Summary

The first two of these initiatives are addressing the generic need for a method of transporting documents over the net. RosettaNet developed it's own method of doing the same thing for a vertical industry as have several other protocols.

Architecture/discovery standards

Before you can carry out eCommerce you need to be able to understand how to interact with the other party.

One of the major initiatives in this area is the eCo Framework (see <http://www.commercenet.com/eco/>). This allows businesses to define:

- the marketplaces they are members of
- the types of product/services they provide
- the eCommerce technology they use

The results can be used to discover how businesses can connect.

ebXML also has architecture discovery standards within the scope.

Registries and repositories

In order to understand XML documents, you need to be able to discover their definition so that the structure of the document can be checked for correctness.

However where do you get the document definition from?

Two initiatives that are addressing this are:

XML.org. This is a repository that holds definitions of XML documents and the codes they use and is run by OASIS. The format of the documents can be either DTDs or schemas

Microsoft's BizTalk.org. This repository also holds definitions of XML documents. However they must conform to Microsoft's XDR schema language.

Summary

To summarize:

common document structures are needed if mapping between the different structures that are will be required to work

a reliable "postal system" is required so that you can be confident that the documents you need will get delivered.

trading communities such as Commerce One's MarketSite can help solve the problems by allowing businesses to connect in a single way yet communicate with a wide variety of different trading partners.

Bibliography

- [HTTP] HTTP is the Hyper Text Transport Protocol. It is the most popular underlying method of transporting data over the web, for example web pages. Most web addresses begin with <http://www....> that indicates that the Http transport should be used
- [RosettaNet] RosettaNet (see <http://www.rosettanet.org>) is a consortium of computer and electronic component manufacturers that has developed standard document choreographies and XML document definitions for carrying out eCommerce within their vertical industry.
- [Schema Language] XML Schema Languages are methods of defining the structure and content of XML documents that permits extensive automatic validation of an XML document by checking it against its schema definition. Schema languages have been developed by Commerce One (SOX) and Microsoft (XDR). At the time of writing (March 2000), both Commerce One and Microsoft are working with the W3C to develop XSDL which will be **the** standard schema language for XML.
- [SMTP] SMTP is the Simple Mail Transport Protocol, and is used specifically for transporting emails over the web.
- [xCBL] xCBL (XML Common Business Library) is a library of XML definitions that was developed by Commerce One by careful analysis of the semantics or meaning behind equivalent EDI transactions. It consists of definitions of XML fragments such as Name & Address, line item and other definitions. Its use is completely open and royalty free. See <http://www.comercone.com/xml>
- [XCC] XCC is software developed by Commerce One that can be used to integrate ERP systems with the web. It allows developers or systems integrators to build XML documents and then wrap them in an envelope for sending documents reliably over the Internet.

Author

David Burdett

Commerce One

Postal Address:

4440 Rosewood Drive

94588 Pleasanton

CA

USA

Telephon: +1 (925) 520 4422

E-mail: david.burdett@commerceone.com

Web: www.commerceone.com

David Burdett - David Burdett is a member of Commerce One's Document Engineering group based in Pleasanton, California. Before joining Commerce One, David was Development Director, Internet, for Mondex International, a subsidiary of MasterCard. David is the author of version 1.0 of the XML based Internet Open Trading Protocols and chaired the development of the OTP Consortium prior to the submission of IOTP to the IETF. David is an experienced Management Consultant and has worked for "big 6" consultancies including Ernst and Young for seven years. He has developed innovative solutions to business problems using a wide variety of technologies including: distributed database, data mining, IS/IT strategy, systems development methods, and electronic commerce, in a wide variety of industries including, finance/banking, oil, healthcare, insurance, and broadcasting. David has also published several articles and presented at seminars internationally on electronic commerce addressing topics such as reliable messaging and barriers to implementation of new software and hardware based payment methods.