How XML Enables Internet Trading Communities and Marketplaces

Dr. Robert J. Glushko
Director, Advanced Technology
Commerce One
2440 W. El Camino Real, Suite 710
Mt. View, CA US
94040
glushko@commerceone.com

Bob Glushko is responsible for XML strategy, standards, and "community applications" at Commerce One, which he joined in January 1999 when it acquired Veo Systems, which he'd helped found in 1997. Prior to Veo Systems, Glushko was the Chief Scientist and Vice President of Passage Systems, a consulting and systems integration firm specializing in SGML-based publishing (which he co-founded in 1992). He has an undergraduate degree from Stanford, a MS (Software Engineering) from the Wang Institute, and a PhD (cognitive psychology) from UC San Diego.

www.commerceone.com

Abstract

This paper explains why XML is rapidly becoming the enabling technology for Internet markets and trading communities. It revisits EDI's basic premise that it is easier to interconnect business systems using "document-based coupling" than in terms of application interfaces and shows how XML can breathe new life into this philosophy. It recognizes the value of preserving EDI's years of experience in designing messages that meet business process requirements and analyzes the technical limitations in both EDI and XML that make the transformation from EDI to XML challenging.

But the ease with which anyone can invent new XML models for particular industries or subject areas is both a primary attraction and a significant threat to the interoperability of messages within and between trading communities. This paper reviews efforts to create standards for XML applications, emphasizing those that consciously strive for a balanced perspective that recognizes the need for EDI and XML to interoperate.

Finally, the paper introduces the challenges posed by the need for documents to be customized for a particular trading community while still being understood and interoperable with documents in other communities. The paper briefly explains how a Common Business Library encoded in an XML schema language is used in the Marketsite™ Marketplace Platform to meet these challenges.

Internet trading communities and marketplaces -- Introduction

An exciting vision of business to business Internet commerce is that of open trading communities, marketplaces, or "virtual enterprises" in which buyers and suppliers of goods and services discover each other, exchange information, conduct transactions, and seize dramatic economic benefits that would be unachievable in the "bricks and mortar" world. Reduced costs, increased revenues, shorter cycle times, lower inventories, and more timely and comprehensive information about customers and business operations are among the many possible payoffs when companies exploit Internet technology for electronic commerce.

The essential benefit of open trading communities or marketplaces is that they offer buyers the largest set of possible suppliers, each of whom has the largest possible market. Each relationship between a supplier's catalog and "back end" processing system and a buyer's purchasing application no longer requires a point-to-point custom integration and yet another document format. Instead, once a company joins an community, its requests for quotes, catalogs and services are potentially available to every other participant, with no incremental integration cost to itself as new companies join, regardless of the buying or selling application each uses.

The goal of creating marketplaces or virtual enterprises by interconnecting business systems is not new. Ideally, companies could conduct electronic commerce in a completely ad hoc fashion, without prior agreement of any kind, and proposals for "Open EDI" and "Plug and Play Commerce" on the Internet predate the XML groundswell of the past few years. But prior to XML, the technology foundations for this vision of electronic commerce simply weren't capable of making it happen.

From "Two at a Time" to Trading Communities

The Promise and Problems of EDI

EDI's basic premise is that it is easier to interconnect business systems and services in terms of the documents they exchange – on which they largely agree to begin with – than in terms of their application interfaces, which inevitably differ. Document-based coupling is looser than coupling via APIs, but loose coupling is better than no coupling at all, as those who have tried and failed to build distributed applications using CORBA or DCOM know.

A major step toward the creation of electronic trading communities seemingly took place when the emergence of EDI Value Added Networks (VANs) as intermediaries eliminated much of the pairwise tyranny of traditional EDI by which big companies had historically imposed proprietary message formats on small ones. VANs offered a great advantage over pairwise connections between suppliers and buyers, which were completely impractical when large buyers required specific document formats as a condition of getting business. With VANs, the supplier makes one connection to the VAN and the VAN does the format translation work of ensuring that the buyer gets information in the desired format.

However, many of the business models now being proposed for Internet trading communities and marketplaces depend on rapid low-cost integration and incremental

evolution of trading partnerships and are still not feasible with EDI, even with VANs in the middle. EDI's syntax is simply less user and programmer friendly than XML's, so significant application-specific development is still required in EDI integrations to do what XML parsers can handle generically.

XML as the Foundation Technology for Trading Communities

Because it is viewed by many as a "a smarter HTML," XML is heading toward HTML's ubiquity while overcoming HTML's inability to encode content in meaningful ways. At the same time, XML is exploiting the twenty year old premise of EDI to focus on the documents that businesses exchange to request and perform services while sidestepping some of the limitations deriving from EDI's syntactic rigidity. Major database, ERP, and "sell-side" commerce software vendors have developed interfaces that let businesses easily expose information to trading partners as XML. As XML support becomes ubiquitous the cost of application integration is predicted to decline dramatically.

Trading Community = Shared XML Document Types

The essence of using XML to implement a trading community or marketplace is for a "market operator" or "market maker" to define the “community standards” for business documents and the protocols for exchanging and routing messages within the community. Then, buyers, suppliers, or other service providers like shippers or payment acquirers can participate if they can produce and consume those documents.

The core idea is that shared document definitions provide an intuitive framework for specifying the business logic and computations that take place on each end of a document exchange. For example, five shared document definitions are implied in these two business rules:

- if you send me a **request** for a catalog, I will send you a **catalog**
- if you send me a **purchase order** and I can fulfill it, I will send you an **invoice** and a **shipping notice**.

How the documents are produced and what actions result when they are consumed are strictly up to the business at each end of the document exchange. This elevates integration from the system level to the business level. It enables a business to present a clean and stable interface to its business partners despite changes in its internal technology implementation, organization, or processes.

"Describe Once, {Buy, Sell} Anywhere"

Web sellers might initially dread that XML-encoded information makes it too easy for buyers or competitors to compare prices, a much wider customer base and access to marketplace services are powerful attractions. Furthermore, many buyers, especially in business-to-business markets, consider price a secondary concern to availability, post-sales service, and other factors.

Transparent Scalability

Defining interfaces in terms of XML documents also allows for an incremental path to business automation, whereby browser-based tasks are gradually transferred to

computer processes. A supplier with a small product catalog and a few sales a day can use a web browser to receive orders and send acknowledgments until increased transaction volume justifies integration with ERP or database applications. Likewise, a buyer who buys only a few items "off the shelf" can rely on a browser to send orders and receive acknowledgments, and only integrate with purchasing or accounting systems when scale justifies it. In each case, since the same XML documents are going in and out, the changes to the implementation are invisible to the marketplace and other trading partners.

XML {vs, and} EDI

XML has rapidly become the first choice for defining data interchange formats in new electronic commerce applications on the Internet and is an overly hyped topic in electronic commerce publications and conferences. Many have interpreted these as signs that "EDI is dead" -- made completely obsolete by the XML upstart -- but this view is naïve from both technical and business standpoints. Companies with large investments in EDI integration will not abandon them without good reason, and if and when they decide to take advantage of new capabilities offered by XML, they will try to preserve as much of those investments as they can. However, transforming EDI to XML is not straightforward and some hard technical problems must be overcome.

Transforming EDI to XML

Twenty years of EDI experience has created X12 [DISA] and UN/EDIFACT [UNEDI] standards for a few hundred transaction sets and messages in many different industries and application areas. So at first glance it might seem that traditional EDI could be quickly adapted to the Internet to obtain lower cost and faster message delivery, while easily enabling integration of EDI-enabled functions with other Web services. In practice, however, EDI will not make the transition to the Internet and XML easily. This section describes four problems that must be dealt with in the transformation from EDI to XML.

The Problem of EDI Subsets

One challenge in adapting EDI to XML is that "standard" EDI messages are never used "as is" in EDI practice. Because the standard messages have evolved through accretion of optional data elements to handle the information requirements of every conceivable business relationship, they contain vastly more information than is typically necessary in any particular case. As a result, the messages that are exchanged between trading partners are always substantially reduced subsets that are heavily customized to that relationship. So if a company exactly preserves its current EDI messages in XML, the more its XML messages will apply only to pre-existing relationships and relatively little will be gained beyond the cost savings in moving from a proprietary message network to the Internet.

Thus EDI has a self-fulfilling bias against the kind of spontaneous commerce to be enabled in open trading communities; because of the historically high cost of EDI integration, companies don't use it unless they have entered into a long term, high volume or high value business arrangement. Once such a point-to-point relationship exists,

though, it is sensible to optimize it by encoding in EDI messages any information that is specific to that relationship, such as contract numbers, buyer catalog numbers, and so on. These optimizations make these messages at worst unintelligible and at best bloated from the perspective of a potential business partner, even in the same trading community.

Some industries that are heavy users of EDI have attempted to combat the proliferation of customized EDI subsets and have developed standard subsets for the most commonly used EDI messages in their trading communities. These subsets are typically called Implementation Guidelines, and using these as the starting point for transformation to XML seems more fruitful than starting from the complete standard messages. The SIMPL-EDI initiative [SIM] seems like a promising place to start efforts to transform EDI standards to XML because its goal is to define messages that "are significantly simpler in content and structure than any previously-published International EDIFACT subset."

The Problem With Programmatic Transformation

Another challenge in transforming EDI to XML is that until very recently EDI has lacked the formal equivalent of XML DTDs for describing the standard messages or Implementation Guidelines in a completely rigorous and computer-processable way. The IMPDEF message in UN/EDIFACT can be used to describe Implementation Guidelines in EDIFACT syntax, and gXML [GXML] is a proposal by an X12 EDI vendor to encode Implementation Guidelines in XML, but neither has been widely adopted so far.

It is hard to expect that standard transformation approaches from EDI to XML can emerge when there aren't any standard starting points. A number of organizations and ad hoc initiatives are working to develop guidelines and tools for transforming EDI data dictionaries and messages into XML versions, but there is little consensus yet about the best technical approach. The X12C /TG3 EDI Architecture Task Group [X12XML] and the CEN/ISSS XML/EDI Workshop [CEN] are two of the most ambitious and careful of these efforts.

The Problem of Different Models of a Message

In addition to business content, EDI messages typically contain the delivery destination and information about workflow or "choreography" such as the message identifier to which the message is a reply, whether acknowledgments are expected, and other information that is needed to deliver the message in accordance with the agreements between the trading parties. Many XML architects would argue that separating the message content from this addressing and workflow information as separate documents and conveying them using multipart MIME would allow better technical approaches for message routing, security, error handling, authentication, confidentiality, and so on. However, no standards yet exist for how XML documents should be "wrapped" with this delivery metainformation and the rules by which it is used to ensure that messages are delivered and processed as they are intended.

The Problem with Modeling Limitations in XML DTDs

A fourth problem with transforming EDI to XML is also a limitation on the XML side of the transform. Unlike EDI, which began with a focus on electronic commerce,

XML has roots in publishing and has come to electronic commerce only afterwards. EDI syntax can thus encode commerce-relevant semantic information about dataypes that XML DTDs, which model element content mostly just as text, can't handle without considerable contortion.

These modeling limitations for XML will be overcome with the XML Schema specification [W3C1, W3C2], to be released by the W3C later this year, which will contain primitive and user-defined datatypes, more expressive occurrence models, and inheritance mechanisms. The richer semantic encoding in XML Schemas will make it easier to preserve the semantics of EDI messages and to transform them into the formats needed by other applications.

XML Standardization Initiatives

The advent of XML Schemas will greatly improve XML's modeling capabilities for electronic commerce but they will do little to address the emerging problem of semantic incompatibility among XML models. Since anyone using XML can invent new models for particular subject areas and define them in a DTD or Schema, what prevents the proliferation of multiple models for the same application or business process? The same content will inevitably be described using different element or attribute names, and different content will be given the same names. XML namespace mechanisms can prevent outright name collisions by prepending a schema name to an element name, but this is solely a syntactic remedy that ignores the issues of semantic incompatibility in the content models with the colliding names.

If every business invents its own XML definitions for product catalogs, requests for quotes, price lists, purchase orders, invoices, transportation schedules, shipping notices, delivery and payment receipts, the Web will become scarcely more usable as a platform for electronic commerce than if everyone used HTML.

Vertical Market or Industry Standards

The solution is industry collaboration and "coopetition" -- many forward-looking individuals and companies have started to work together to develop XML-based specifications for the information they most often need to exchange in a particular industry or vertical market. These XML standardization initiatives like RosettaNet (computer supply chain) [RN], OAG (enterprise application integration) [OAG], OBI (purchasing) [OBI], OTP (payment) [OTP], OTA (travel) [OTA], and so on are critical enablers of markets and trading communities. The vendor-neutral OASIS organization [XMLORG] and the Microsoft-sponsored Biztalk initiative [BIZ] are both developing registries and repositories in which companies and industry groups can make their XML definitions available to others.

Vertical communities can have very rich content and specialized processes, which imply highly specialized document models. Thus there are substantial benefits when XML definitions are shared by the companies in a particular industry or market. These include reduced development and maintenance costs and the elimination of custom "mapping" between the information models embodied in a company's business systems and those of its trading partners.

Horizontal Standards

But while each new XML specification for a particular industry or marketplace is a step forward for that industry, each contributes to a different interoperability problem because they proliferate definitions of information models that cut across industries or marketplaces in which a single company may need to participate. Some concepts and constructs needed in these "vertical" specifications apply to all business domains, but each new specification seems to "start from scratch" and reinvent them. For example:

- Descriptions of businesses and individuals;
- Measurements, date and time, location, country codes, currencies, business classification codes;
- Basic business forms like catalogs, purchase orders, and invoices.

Any large company will sell products in both direct and indirect markets, maintain a supply chain for its direct inputs to its manufacturing processes, procure large amounts of indirect goods for its operations, post job offers in employment marketplaces, and so on. If each of these domains develops its own schemas for the basic documents, it is inevitable that some of them will be incompatible.

It might be funny to say that the "O" in OAG, OTA, OTP, OBI, and so on stood for "overlapping" were the consequences not so serious. It isn't enough that people are developing XML specifications for specific industries and applications. In addition, there needs to be a way to encourage the development of XML document models from reusable semantic components that are common to many business domains. Such documents can be understood from their common message elements, while also providing a common mechanism for linking to unique elements that vendors need to differentiate themselves.

The Common Business Library

The oldest attempt to attack the problem of interoperability among vertical XML commerce applications is Commerce One's Common Business Library [CBL]. CBL proposes a set of reusable XML components that are common to many business domains, along with a set of document frameworks for creating documents with a common architecture. Documents built according to the CBL frameworks can be understood from their common message elements and extended in predictable ways.

Work on CBL began in 1997, partly funded by a Department of Commerce's Advanced Technology Program research award on "Component-Based Commerce" to Veo Systems and three other firms [ATP]. Because of this research pedigree, early versions of CBL strove for logical completeness, expressiveness, and compactness to test the abstract modeling power of XML for electronic commerce and to identify requirements for development tools and runtime support. CBL 1.0 prototyping and application experience suggested that it was too abstract and powerful for XML “newbies” and for people with traditional EDI backgrounds, both of whom preferred document types more analogous to familiar business forms. Furthermore, XML DTDs and “off the shelf” XML tools weren’t capable of handling the modular reuse and extensibility intended for CBL 1.0, which made extensive use of typed pointers to allow

creation of compound documents in a disciplined way, and liberal use of parameter entities to allow for customization [ALL].

The acquisition of Veo Systems by Commerce One in January 1999 introduced to CBL a requirement for interoperability with EDI. CBL 2.0, in line with the lessons learned from CBL 1.0, aims for less abstraction, even if it means redundancy or less expressiveness, and greater compatibility with EDI standards and semantics. Using standard data element semantics provides a strong non-proprietary and interoperable semantic foundation for CBL, and gives companies using EDI today a clear migration path in CBL for transforming EDI applications to XML. CBL is freely available in repositories run by Commerce One as part of marketsite.net, as well as through those operated by xml.org and Biztalk.org.

The CommerceNet eCo Architecture

A cross-industry specification of a different kind has been proposed by the eCo Working Group chartered by the CommerceNet Consortium [ECO]. The eCo group recognized that the great pace of innovation in electronic commerce architectures and implementations makes it unlikely that a single standard would emerge soon. This means that a business that wants to combine its services with others to create a trading community or marketplace might have to deal with an incompatible variety of implementations, protocols, and business processes. This diversity raises the implementation cost and limits the alternatives for companies who want to establish and maintain multiple business relationships.

So rather than attempting to define protocols or document models, the eCo Group proposed a framework for defining "a world in which different ones can co-exist." The eCo architecture specification, published in October 1999, defines a reference model for describing those aspects of electronic commerce systems that are relevant to interoperability. The specification presents XML schemas for describing marketplaces, the businesses that belong to them, the services provided by those businesses, and the document interchanges that implement each service. Thus the eCo specifications could describe EDI implementations that have been moved to the Internet by "wrapping" existing systems and processes with standard eCo interfaces.

The ebXML Initiative

The fundamental problems of designing messages that meet business process requirements and the standardization of their semantics are independent of the syntax in which the messages are encoded. So it is extremely encouraging that the EDI and XML communities are coming together in an initiative whose goal is nothing less than to determine "the technical basis upon which the global implementation of XML can be standardized" for electronic commerce. This initiative, called ebXML (for "Electronic Business XML"), was jointly announced by UN/CEFACT and OASIS in September, 1999 [EBXML]. It will develop global, syntax-independent message design guidelines that harmonize EDI and XML architectures and has the potential to create a standard encoding for transforming and representing EDI semantics in XML.

The ebXML initiative has great promise. It is both the first XML standardization activity begun by a global EDI standards body and the first attempt by the EDI standards

community to work with XML experts as equal partners in shaping the transformation of EDI to XML. If ebXML succeeds in attracting participation from a critical mass of the XML specifications currently proposed or under development, it will speed their convergence to interoperable architectures.

The Challenges of Evolution and Interoperation

But even if ebXML or any of the numerous XML specifications were to be widely adopted as the standard business documents, large challenges remain for XML trading communities. Documents defined for a community will not be fixed; rather they will constantly evolve due to changes in regulations, business processes, service offerings etc. This fast paced change will be the hallmark of future business and its support will be a necessity for future success of electronic commerce systems.

XML DTDs are somewhat limited and inflexible in their support for controlled customization and extensibility. If a DTD must be changed after the fact to allow for unanticipated customizations, all of the applications relying on the original DTD must also be changed, even if they don't need the customizations.

Most XML architects working on Internet markets and trading communities are hopeful that XML Schemas will be the technical magic that enables documents to be customized for a particular trading community while preserving the interoperability with other communities. While the emerging W3C recommendation for XML Schemas is not completely determined as this paper is being written in October, 1999 [W3C1, W3C2], schema extension mechanisms are certain to be included. Schema extensions would allow a community to define base documents that could be extended to meet the customization needs for some trading relationships while maintaining the ability to ignore those extensions in other relationships or communities where they are not needed.

Commerce One's MarketSite™ Marketplace Platform

The software solution today that comes closest to realizing the potential of XML to enable Internet marketplaces and trading communities is Commerce One's Marketsite™ Marketplace Platform [COM]. The Marketsite Marketplace Platform architecture provides a single integration point for suppliers, buyers, and commerce service providers. The Platform is being used by Commerce One to host the www.marketsite.net business to business procurement community and for trading communities operated by British Telecom, NTT, Singapore Telecom, Cable and Wireless Optus, PeopleSoft, Warner-Lambert, Schlumberger, and many other Commerce One customers.

By making these regional and vertical trading communities work together, Commerce One is forming a "Global Trading Web" in which businesses of all sizes can enable their employees to source, buy and sell goods and services on a global basis, in real time [GTW]. The Global Trading Web can be accessed by a wide range of commercially available buying and selling applications, including Commerce One BuySite, as well as applications from Intershop, PeopleSoft, RightWorks and SAP.

The interoperability among all of these Internet marketplaces and the buying and supplier systems that connect into them is enabled through the use of application

interfaces defined using the reusable components of the Common Business Library encoded in XML schema language called the Schema for Object-Oriented XML [SOX]. SOX, which has been submitted to the W3C XML Schema Working Group, supports the XML namespace mechanisms for building schemas from previously defined element types, so trading communities can define documents that have elements in common with other CBL-based communities or from other XML electronic commerce specifications. This greatly enables businesses to participate in multiple marketplaces with the same documents, minimizing document transformation and integration headaches.

Furthermore, because SOX has mechanisms for extending element content models, the standard documents used by the community can be customized to add information needed by particular trading partners or for a business to enter a new market. Applications relying on the base schema would not need to be changed to handle instances of the extended one. See [KOS] for an example of how SOX's extension mechanisms enable a construct like "Shipping Address" to be based on a standard and simpler "Address" model.

Commerce One was a corporate sponsor and active participant in the CommerceNet eCo working group. The eCo architecture specifications will be supported in the Marketsite Marketplace Platform to ensure that it can interoperate with other Internet marketplaces -- not just those built using Commerce One's software, further growing the Global Trading Web.

References

[ALL] Allen, T. *Common Business Library (CBL).* Proceedings of XML Europe '99.

[ATP] *Component-Based Commerce: The Interoperable Future*, http://jazz.nist.gov/atpcf/prjbriefs/prjbrief.cfm?ProjectNumber=97-06-0032

[BIZ] BizTalk.org, www.biztalk.org

[CBL] Common Business Library, http://www.marketsite.net/xml/aboutcbl.html

[CEN] European Committee for Standardization/Information Society Standardization System (CEN/ISSS) XML/EDI Workshop www.cenorm.be/ISSS/Workshop/XMLEDI

[COM] Commerce One MarketSite Open Marketplace Platform, http://www.commerceone.com/solutions/marketsite.htm

[DISA] Data Interchange Standards Association, ASC X12 Interindustry Standards for the Electronic Interchange of Business Transactions, www.disa.org

[EBXML] *United Nations and OASIS Join Forces to Produce Global XML Framework*, 15 September 1999, www.oasis-open.org/html/OASIS_UN.html

[ECO] The eCo Framework. CommerceNet Consortium, www.eco.commerce.net

[GTW] *Commerce One Redefines Electronic Commerce with World's Largest, Open Business-to-Business Marketplace*, 29 March 1999, http://www.commerceone.com/news/US/openmarketplace.htm

[GXML] *gXML - Open Exchange of E-business Schemas*, Edifecs Commerce, 30 March 1999, www.edifecs.com/content/corporate/pressrelease03309900.asp

[KOI] Koistinen, J., Davidson, A., Fuchs, M., Jain, M., and Schwarzhoff, K. *XML Programming Models for Electronic Commerce Systems*. Proceedings of XML'99 Conference, December 1999.

[OAG] Open Applications Group, http://www.openapplications.org/

[OBI] Open Buying on the Internet, http://www.openbuy.org/

[OTA] Open Travel Alliance, http://www.disa.org/opentravel.com/index.htm

[OTP] Internet Open Trading Protocol, http://www.otp.org/

[RN] RosettaNet, http://www.rosettanet.org/

[SIM] SIMPL-EDI, Association for Standards and Practices in Electronic Trade - EAN UK Ltd, www.ecentre.org.uk/products_businesstech_simpledi.asp

[SOX] Davidson, A., Fuchs, M. Hedin, M., Jain, M., Koistinen, J., Lloyd, C., Maloney, M., and Schwarzhoff, K. *Schema for Object-Oriented XML 2.0*, Commerce One, July 1999. www.w3.org/TR/NOTE-SOX

[UNEDI] UNITED NATIONS DIRECTORIES FOR ELECTRONIC DATA INTERCHANGE FOR ADMINISTRATION, COMMERCE AND TRANSPORT, www.unece.org/trade/untdid/welcome.htm

[W3C1] XML Schema Part 1: Structures, W3C Working Draft 24 September 1999, http://www.w3.org/TR/xmlschema-1/

[W3C2] XML Schema Part 2: Datatypes, W3C Working Draft 24 September 1999, http://www.w3.org/TR/xmlschema-2/

[X12XML] The X12C /TG3 EDI Architecture Task Group (http://ww.disa.org/x12/x12c/X12CTG3/PDF/xmltechreport.pdf)

[XMLORG] XML.org, www.xml.org