

April 1999

# XML and Its Potential for Ecommerce

Norman Desmarais

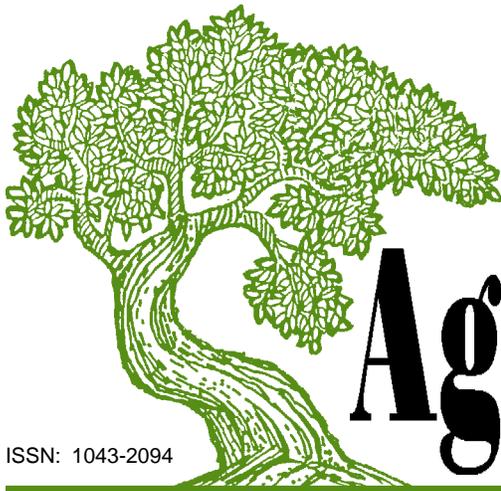
*Providence College*, normd@providence.edu

Follow this and additional works at: [http://digitalcommons.providence.edu/facstaff\\_pubs](http://digitalcommons.providence.edu/facstaff_pubs)

---

Desmarais, Norman, "XML and Its Potential for Ecommerce" (1999). *Library Faculty and Staff papers*. 9.  
[http://digitalcommons.providence.edu/facstaff\\_pubs/9](http://digitalcommons.providence.edu/facstaff_pubs/9)

This Article is brought to you for free and open access by the Phillips Memorial Library at DigitalCommons@Providence. It has been accepted for inclusion in Library Faculty and Staff papers by an authorized administrator of DigitalCommons@Providence. For more information, please contact mcaprio1@providence.edu, hposey@providence.edu.



# Against the Grain

TM

*"Linking Publishers, Vendors and Librarians"*

ISSN: 1043-2094

## Chaos — XML and Its Potential for ECommerce

by **Norman Desmarais** (Providence College) <normd@sequent1.providence.edu>

Column Editors: **Sandra K. Paul** (President, SKP Associates) <SKPAssoc@cwixmail.com>  
and **Albert W. Simmonds** (SKP Associates) <awsimmo@ibm.net>

Electronic commerce did not emerge with the development of the World Wide Web or the commercialization of the Internet. It has been around for over 25 years. It germinated in 1948 during the Berlin airlift when **Ed Guilbert** of the **Department of Defense** spearheaded the **Transportation Data Coordinating Committee**. This group needed to get the various modes of transportation (air, railroad, and ocean) to share data like schedules and to coordinate with each other and with their clients to effectively deliver goods and supplies. This eventually led to the establishment of the **X12** committee which was accredited by **ANSI (American National Standards Institute)** in 1979. Large petroleum, banking, transportation, and retailing companies and the federal government soon began using **X12 EDI (Electronic Data Interchange)** to provide electronic forms and messages for shipping and purchasing. However, the book industry had already begun establishing communication standards in 1974.

It is easy to equate **Electronic Data Interchange** with electronic commerce, but **EDI** is really only one element of electronic commerce. We can describe Electronic Data Interchange as the communication of business information using "a collection of standard message formats and element dictionaries to exchange data." It is a set of standards that facilitate business to business processes such as ordering and fulfillment and financial transactions. Electronic commerce, on the other hand, could be defined simply as "doing business electronically." The concept implies the exchange of information over computer networks; but it also involves online querying of supplier databases and the realtime integration of supply chains over the Internet and extranets. Electronic commerce could use structured communications (such as EDI) or unstructured messages (such as email) to transmit information, data, databases, data-

base access, or a combination of these. Electronic commerce comprises at least three types of communications: consumer to consumer, consumer to business, and business to business and it implies electronic links between the various sources of information.

We shall begin by reviewing the origin and development of standards for Electronic Data Interchange, considering exceptions and variations that standards need to accommodate, and examining cost factors that affect implementation decisions. We shall then proceed to discuss conducting EDI over the Internet and the importance of metatags and data type definitions for interpreting data structures and improving functionality of business transactions. Then, we'll look at **XML (eXtensible Markup Language)** as a proposal to broaden the applicability of transacting business over the Internet by bringing EDI to the desktop. We shall consider how XML could work for EDI over the Internet and the benefits that it could affect.

### Origin and Development of EDI Standards

Although humans have long envisioned that computers would improve communication and data interchange for commercial applications, reality usually required developing custom interfaces for each different system to facilitate transactions with suppliers and customers. In the six years from its creation, in 1974, until the **Book Industry Systems Advisory Committee (BISAC)** became a committee of the **Book Industry Study Group (BISG)** in June, 1980,

BISAC aimed to streamline the ordering and supply of books and printed materials. It developed formats for purchase orders, order acknowledgment, invoice, title status, payment advice, frontlist diskette, data transmission protocols, royalty statement, and sales reporting.

In 1979, the **American National Standards Institute (ANSI) Accredited Standards Committee (ASC) X12** began to develop uniform standards for the electronic exchange of business transactions in standards known as electronic data interchange (EDI). Around 1983, the major BISAC members wanted BISAC to develop formats based on the X12 formats to allow the book industry to follow standards comparable to other industries and supported by standardized "translators" and value added networks.

The ASC X12 committee meets three times a year to develop and maintain EDI standards. The committee's main objective is to "develop standards to facilitate electronic interchange relating to such business transactions as order placement and processing, shipping and receiving information, invoicing, and payment and cash application data, and data to and from entities involved in general business, finance, insurance, education, and state and federal governments."

The **ASC X12** standards aim to facilitate electronic commercial transactions by establishing a common, uniform business language for computers to communicate. They comprise more than 300 transaction sets that allow businesses to execute nearly every type of business-to-busi-

*continued on page 87*



ness operation electronically, such as orders, invoices, customs declarations, statistics, insurance documents, bills of transport, and many health care transactions. They are often the electronic equivalent of the preprinted forms sold in stationery stores. The goal was not to replace paper forms with electronic equivalents, but to allow businesses the flexibility to reengineer how work is processed in order to gain efficiencies of working electronically via EDI. Those companies that closely examined their work processes benefited more than those who just replaced paper forms with electronic forms.

The EDI standards have even become recognized internationally and have received endorsement by the **United Nations** which developed the **UN/EDIFACT** standard (**United Nations Standard Messages Directory for Electronic Data Interchange for Administration, Commerce, and Transportation**), in many respects, based on lessons learned in X12's development experiences. **UN/EDIFACT** messages use predefined field identifiers which must occur in a predefined sequence. Some users claim it is less robust than **ANSI ASC X12**. Others say the opposite. But its widespread acceptance makes it viable. EDI got further endorsement in 1997, when **President Clinton** decreed that all U.S. government procurement be accomplished via electronic commerce. The U.S. government then began moving toward adopting and implementing **X12** and **UN/EDIFACT**.

#### **Exceptions and Variations**

Every industry has its own way of doing things. Not only must international standards accommodate possible variations by industry and company, they must also take into consideration each country's own set of exceptions. Even within the same industry, many companies operate differently. It is not uncommon to add comments on a purchase order to clarify what one wants or the terms of delivery, for example. Standard EDI messages must support these comments and many other options that make them very complex. This complexity, resulting from the differences between trading partners and the related costs of programming or mapping data, discourage many companies from implementing EDI.

Also, in the development of international EDI standards, language becomes important. We take it for granted that the data elements and syntax use English. English is a dominant language in the business world and the Internet is based on English syntax and almost every program is coded in English language friendly computer languages. However, have French or Chinese vendors and buyers, for example, been willing to use English? Should the data elements be language neutral or should multiple naming conventions and syntaxes be used?

#### **Cost Factors**

EDI messages serve to exchange documents or data between companies. However, different organizations store their data in different formats and representations (data layouts). One

company may store its data in a relational database, for example, and may have to map it differently for transmission via EDI. Such mapping can be very involved and costly.

In addition to the costs of programming and data mapping, companies incur high costs in acquiring and implementing EDI. These costs include hardware, setup, EDI enabling software, transaction service fees, telecommunications charges, and annual software maintenance contracts that the ongoing evolution of EDI standards require.

Companies must realize sizable savings to break even. Because the quantity of documents exchanged determines the amount of savings, large companies that process many transactions can offset their costs in a relatively short time. Small and medium-sized organizations, on the other hand, will find it harder to justify EDI. Large companies can offset the cost of EDI through increased efficiency. Smaller companies, however, can find the costs prohibitive, excluding them from revenue opportunities. For example, some large companies, particularly those in low-margin retail businesses such as **WalMart** and **Kmart**, refuse to conduct business with any company that does not use EDI.

The wide range of interfaces employed by users further complicates implementation and support. In addition to creating, or contracting with a software vendor to create a custom interface for every trading partner with which it exchanges data, a company must deal with the complexity and expense of custom user interfaces that must be created for each new form, document, or process. This makes EDI a complicated and expensive proposition for many.

#### **EDI Over the Internet**

EDI has been most successful in vertical markets and in business-to-business transactions with a large number of interactions. The advent of the Internet and its use for business applications changed how people interact when they exchange goods and services. EDI is no longer restricted to business to business communications. The Internet has introduced many new ways of trading, allowing interaction between groups that previously could not afford to trade with one another economically. The same principles apply to all commercial transactions, whether the consumer is an end user, a manufacturer, a service organization, a governmental agency, or a virtual organization. So all participants in the electronic marketplace should be able to transact business with the same ease.

The **Internet Engineering Task Force (IETF)**, a large open international community of network designers, operators, vendors, and researchers concerned with the evolution of the Internet architecture and the smooth operation of the Internet, is part of **The Internet Society** that develops Internet standards. Much of its work is done in its working groups. Some of these groups are concerned with business com-

munications and should be making recommendations for implementing standards of commerce over the Internet.

A major challenge to performing EDI on the Internet is to reduce the cost of doing EDI enough to allow small and medium-sized organizations to realize savings also. The goal is to enable them to perform EDI transactions with only a browser and an Internet connection. The application of **eXtensible Markup Language (XML)** to EDI on the Internet promises to do just that.

#### **Metatags and Data Type Definitions**

By using metatags, such as those employed in XML, to identify data elements, companies can tag their data once and use it with any number of applications, such as in the preparation of catalogs, purchase orders, and invoices. The

tags also allow for distinguishing different meanings for the same word. For example, one could use different tags to identify the term "pocketbook" as a purse, a billfold or wallet, or a small book. The term "date" with no qualifiers could indicate an order date, a shipping date, a received date, a social engagement or a person's companion on such an outing, or even a fruit.

XML tags and attributes can incorporate long-established messaging structures and rules that vary from industry to industry into Web-based documents for Electronic Data Interchange between trading partners. This could streamline the purchase of goods and open the supply chain to any trading partner with an Internet connection and an XML-capable browser. It could also facilitate the entry of smaller companies into the electronic marketplace by allowing vendors to offer products directly to a buyer, effectively removing the middleman because all processing is done on the client side, the buyer's computer. Buyers and suppliers of all sizes can participate in commerce over the Internet because it does not require the use of proprietary technologies.

As the **ANSI ASC X12** and **UN/EDIFACT** standards have already devised structured messages for most types of business transactions, it's very likely that they can serve as the basis for **XML Data Type Definitions (DTD)** for commercial transactions. The DTD and metatags will be most useful when used by a number of people and tools. However, in the real world, people often agree to comply with the standards, then modify them a little or a lot to suit their own needs and those with whom they're communicating. While XML offers a high degree of flexibility, industry users will need to resolve how to establish data dictionaries and how to make those dictionaries usable to general purpose electronic commerce applications. **Brian Green, Secretary of the European Book Sector EDI Group**, expects that **EDItEUR** will soon begin developing data neutral data dictionaries which will allow **UN/EDIFACT** tags to

*continued on page 88*



be used by XML or any other language that uses metatags.

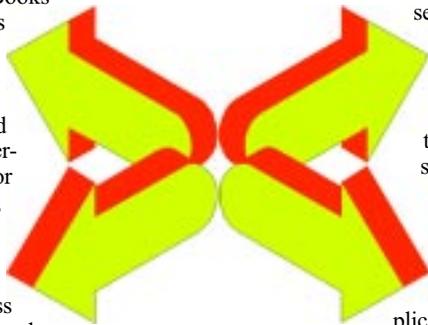
Part of the problem of implementing EDI in the library market is the very complex product description. The **MARC** format could serve as the DTD for library applications as it contains all the information needed by librarians and anybody working in the book trade. However, any changes will likely occur over a long period of time, and business requirements will determine how this will all evolve.

A DTD could serve to specify the structure of a document and how to interpret it, much like a template. The creation of a template in this fashion could enable users to create the equivalent of EDI messages without the labor and cost of extensive data mapping. Librarians process an enormous number of transactions; but the relatively low price per item and small margins currently militate against widespread implementation of EDI. By implementing XML, however, librarians and vendors can process messages on the client side, further reducing costs and processing overhead.

#### Metatags Mean Greater Functionality

Today's search engines can easily extract information from textual databases and Web pages. But some form of tagging system must be used to retrieve "fielded" data accurately. XML/EDI and the use of metatags permit using structured data along with unstructured text in the same document. These metatags allow searching, decoding, manipulating, and displaying data consistently and correctly without the need to create special interfaces. They could also enhance commerce on the Web by adding new dimensions not currently available. For example, if online booksellers use a standard set of metatags to mark titles, descriptions, and prices, buyers could quickly search all of the online bookstores for the lowest price on a book with a single command. Books

and music have emerged as the leading goods and services sold on the Internet. **Zona Research Inc.** (Redwood City, CA) estimated that book sales over the Internet totaled \$181 million for the fourth quarter of 1997, with music trailing at \$47 million.<sup>1</sup> It forecasts that the number of businesses planning to conduct business over the Internet will grow by 34%. **International Data Corp.** (Framingham, MA) estimates that the amount of commerce conducted over the Internet will increase from \$2.6 billion in 1996 to \$220 billion in 2001.<sup>2</sup> The market is so lucrative that **Barnes & Noble** is reportedly prepared to lose \$7 million dollars a year to build market share online.<sup>3</sup> **Robert Krulwich**, in an **ABC News** broadcast on November 24, 1998, reported that **Barnes & Noble's** customer base stood at 930,000 compared to **Amazon's** 4,500,000. While **Amazon's** new customers increased by 1,200,000 in the last few months, **Barnes & Noble** only added 210,000.



However, this was before the purchase of half of **Barnesandnoble.com** by **Bertelsmann**.

Online catalogs are the first applications taking advantage of XML and major European wholesalers publish monthly multimedia CD-ROM stock catalogs. Most of these catalogs allow users to build order files and to transmit them in EDI formats, normally using direct dialup. But the number of consumers who purchase goods over the Internet account for only one percent of the consumer market.

Today, only sites that have been specifically set up to exchange information can conduct EDI transactions. XML would permit exchanging data regardless of the computing systems or accounting applications being used. The intent of XML/EDI is to establish a standard for unambiguous commercial electronic data interchange that is open and accessible to all and which can satisfy the full breadth of business needs. **Marty Tenenbaum**, chairman of **CommerceNet**, says that "it's essential that all these systems talk to each other; and they can't today, except at the level of HTML."

#### Planning for the Future

Attaining these objectives requires using a scalable or extensible means that will satisfy current needs as well as future requirements, including the ability to incorporate new technologies and business needs as they emerge. The selected technology also needs to be widely and freely available to ensure widespread adoption. XML has been proposed as a solution that meets these requirements.

The **W3C** published a specification for a formal **Document Object Model (DOM)** for XML documents in October, 1997. This model provides a standardized **API (application programming interface)** for XML-based tools. It indicates that XML-coded electronic forms may become the main method of capturing and coding EDI information.

#### Using XML for Electronic Data Interchange

The use of XML for the interchange of commercial EDI messages follows several stages. First, one must identify the suitable data sets to use in electronic business transactions, such as the ones in the UN/EDIFACT standards. Then, one develops the XML data type definitions (DTDs) that specify how the fields that are to form a particular class of EDI messages relate to one another. Third, come the definitions of the relationships of the fields and of application-specific extensions to standard message types. Then, one creates the specific types of electronic business messages, validates the contents of messages, and transmits and receives electronic business messages. Finally, one could use **DataBots**, data manipulation agents, to process electronic business messages.

Because **XML/EDI** is a standard for formatting documents, it leaves the method for how to store or transmit these documents to the underlying applications. One can use FTP, Email (SMTP), HTTP, or another method to transmit XML/EDI transactions just as any other document. One can

transmit them via a value added network (VAN), Intranet (WAN), or Internet. One could also store them on diskette or other transportable medium.

XML could use the XML protocol as its "data interchange modeling" layer and the **XSL (Extensible Stylesheet Language)** protocol as its "presentation" layer. XML would allow the data to interface with EDI, but the underlying processing would remain EDI. XML could function with traditional methods of EDI and could support all standard Internet transport mechanisms such as IP routing, HTTP, FTP, and SMTP. It uses programming tools such as **Java** and **ActiveX** to permit data sharing between programs; and it uses agent technologies to manipulate, parse, map, and search for data.

#### Benefits

Leading the list of benefits of using XML for e-commerce is that it allows people and companies to exchange information more clearly and completely than previous formats. It lets users conduct business transactions at the client level on desktop computers with a general purpose document or Web browser as the user interface. A common language describes the templates and associated conversion rules. Even though XML/EDI allows distributed processing capabilities, it also supports centralized functions. It lets users be more independent, permitting them to define and issue documents rather than relying on data processing specialists to do so.

Instead of having to create templates and interfaces for each trading partner, traditional EDI service providers will just need to interface their products to XML/EDI templates. As companies implement XML/EDI to provide simpler to use types of transaction tools, businesses that fail to incorporate XML/EDI into their operations will limit their growth and may see their products eliminated from the marketplace because their competition will have made it easier to do business with them.

Because XML/EDI supports legacy or traditional EDI systems by definition, companies that use UN/EDIFACT or ASC X12 standards can continue to use their systems. They can create a conversion routine (gateway) between the subsystems that use UN/EDIFACT or ASC X12 and those that use XML/EDI to interface with clients or suppliers who are reluctant to adopt XML/EDI. Business requirements will determine how this will all evolve and will likely occur over a long period of time.

Some businesses are starting to look at XML as a means to reduce procurement costs when buying goods over the Internet. **Commerce One** has a product, **BuySite**, that automates the procurement process from requisition to order. Its **MarketSite** software then automates the interaction with the supplier from order placement to payment. The company expects that integrating XML will reduce operational costs and increase efficiency because of its ability to define products so buyers can search the online catalogs to identify the right product at the right price.

We briefly reviewed the history of the development of standards for Electronic Data Interchange, considered some of the exceptions and variations that standards need to account for, and examined cost factors that affect decisions to implement EDI. We discussed how the

*continued on page 89*

**Chaos**  
from page 88

Internet is changing expectations for electronic business transactions; and we considered XML as an emerging solution. By applying object-oriented technology to EDI, XML uses metatags and data type definitions to interpret data structures and improve the functionality of business transactions. XML promises to allow clients to implement EDI with a minimum of effort,

thereby reducing costs. Implementation will not occur overnight. The change will be evolutionary rather than revolutionary. As applications get updated and modified with XML code, these solutions will gradually replace older, more cumbersome applications. In the meantime, several companies and organizations have begun to adapt existing standards or to prepare DTDs for use with XML. Others are developing tools to facilitate the transition. 

**Endnotes**

- <sup>1</sup> LAN Times Online 1/98.
- <sup>2</sup> Hurley, Hanna. "EDI Takes to the Internet." *Network Magazine*. Oct. 1998, p. 36.
- <sup>3</sup> *The Search for Digital Excellence*. James P. Ware et al. New York: McGrawHill, 1998, p. 29.
- <sup>4</sup> Messmer, Ellen. "Software Aims to Bridge EDI/XML." *Traffic. Network World*. July 13, 1998, p. 6.

**Bibliography**

Boeri, Robert J. & Martin Hensel. *XML: The New Document Standard*. Emedia Professional. June, 1998.

Dalton, Gregory. XML On The Rise: Language expected to aid Ecommerce, help Web searches. January 26, 1998. <http://techweb.cmp.com/tw/666/66iuxml.htm>.

Drummond, Rik and Kay Spearman. "XML Set to Change the Face of ECommerce." *Network Computing*. May 1, 1998, pp. 140-144.

Goldfarb, Charles F. and Paul Prescod. *The XML Handbook*. Upper Saddle River, NJ: PrenticeHall, 1998.

Gottesman, Ben Z. [et al.]. "Why XML Matters." *PC Magazine*. Oct. 6, 1998, pp. 215238.

Guidelines for using XML for Electronic Data Interchange. Version 0.05. Martin Bryan, Editor. 25th January 1998 <http://www.geocities.com/WallStreet/Floor/5815/guide.htm>.

Heid, Jim. "XML Spoken Here: New Web Language Promises Smarter Surfing." *PC World*. August, 1998, p. 70.

Hurley, Hanna. "EDI Takes to the Internet." *Network Magazine*. Oct. 1998, pp. 36-40.

Kerstetter, Jim. "XML Holds Promise as EDI Replacement." *PC Week*. May 1, 1998.

Klinger, Steven P. "XML Expands Web App Possibilities." *LANTimes*. June 8, 1998, p. 26.

Krantz, Michael. "Keeping tabs online: doing business on the Net is hard because the underlying software is so dumb; XML will fix that. (new markup language)." *Time* 150: 20 (Nov 10, 1997), pp. 81.

Leventhal, Michael; David Lewis; and Matthew Fuchs. *Designing XML Internet Applications*. Upper Saddle River, NJ: PrenticeHall, 1998.

Levitt, Jason. "The Making of a Markup Language." *Information Week*. May 25, 1998 pp. 70-80.

Light, Richard. *Presenting XML*. Indianapolis: Samsnet, 1997.

Marchal, Benoît. Electronic Data Interchange on The Internet. [http://developer.netscape.com/viewsource/marchal\\_edata.htm](http://developer.netscape.com/viewsource/marchal_edata.htm).

*Markup and Metadata*. Digital Publishing Technologies. April, 1998, pp. 11-16.

McGrath, Sean. *XML by Example: Building ECommerce Applications*. Upper Saddle River, NJ: PrenticeHall, 1998.

Megginson, David. *Structuring XML Documents*. Upper Saddle River, NJ: PrenticeHall, 1998.

Messmer, Ellen. "Software Aims to Bridge EDI/XML." *Traffic. Network World*. July 13, 1998, p. 6.

Millman, Howard. "A Brief History of EDI." *Infoworld* April 6, 1998, p. 83.

Simpson, John C. *Just XML*. Upper Saddle River, NJ: PrenticeHall, 1998.

St. Laurent. Simon. *XML: A Primer*. Foster City, CA: MIS Press, 1998.

Stanek, William Robert. "PC Labs Reviews." *PC Magazine*. October 6, 1998 <http://www.zdnet.com/pcmag/features/xml98/xml201.html>.

Stanek, William Robert. "Structuring Data with XML Extensible Markup Language lets developers describe virtually any type of data and deliver it across a network." *PC Magazine* 5/26/98. <http://www.zdnet.com/pcmag/pctech/content/17/10/tf1710.001.html>.

Swenson, John. "XML: A better way to move data across the Web." *MSDN Online*. January 21, 1998. <http://www.microsoft.com/msdn/news/imho/012198.htm>.

Udell, Jon. "XML describes structured data packages that move around the Web as easily as HTML." *Technology*. November 10, 1997, vol. 150, no. 20. <http://www.byte.com/art/9801/sec5/art18.htm>.

World Wide Web Consortium. Extensible Markup Language (XML) 8 December 1997 <http://www.geocities.com/WallStreet/Floor/5815/>; <http://www.w3.org/TR/PRxml971208>.

**"Linking  
Publishers, Vendors  
and Librarians"**



**Against the Grain**

**Uncommon ...**

*Against the Grain* is your key to the latest news about libraries, publishers, book jobbers, and subscription agents. **ATG** is a unique collection of reports on the issues, literature, and people that impact the world of books, journals, and electronic information.



**Against the Grain**

**Unconventional ...**

**ATG** is published six times a year, in February, April, June, September, November, and December/January. A six issue subscription is available for only \$40 U.S. (\$50 Canada, \$70 foreign), making it an uncommonly good buy for all that it covers. Make checks payable to **Against the Grain, LLC** and mail to:

**Katina Strauch**  
209 Richardson Avenue  
MSC 98, The Citadel  
Charleston, SC 29409

Name	Address	City	State	Zip
			Phone	
			Company	
			Email	