April 1999

Chaos-XML and Its Potential for ECommerce

Norman Desmarais
Providence College, normd_1@providence.edu

Sandra K. Paul
SKP Associates, SKPAssoc@cwixmail.com

Albert Simmonds
SKP Associates, awsimmo@ibm.net

Follow this and additional works at: https://docs.lib.purdue.edu/atg

Part of the Library and Information Science Commons

Recommended Citation
DOI: https://doi.org/10.7771/2380-176X.3852

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
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 Gulbert 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, database 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 afford.

**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-business 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 stationary 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, continued on page 87
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 inter- actions. 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 communications 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 as well. 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.

**continued on page 88**
XML tags and attributes can incorporate long-established messaging structures and rules that vary from industry to industry. This could streamline the purchase of goods and open the supply chain to trading partners. This could also facilitate the entry of smaller companies into the electronic marketplace by allowing them to offer products directly to a buyer, effectively removing the middleman. However, the real world, people work together to resolve problems, and the needs of those with whom they are communicating. 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.

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 required by librarians and anybody working in the library. 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. Libraries can process an enormous number of transactions; but the relatively low price per item and small margins currently mitigate 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-priced book on a single command. Books and music have emerged as the leading goods and services sold on the Internet.

**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, one defines the relationships of the fields and of application-specific extensions to standard message types. Then, one creates the types of electronic business messages, validates the content 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 messages just as any other document. One can transmit them via a value-added network (VAN), Intranet (VAN), or Internet. One could also store them on diskette or other transportable medium.

XML could use the XML protocol as its “data interchange modeling” and the XSL (Extensible Stylesheet Language) protocol as its "data interchange modeling".
to allow the data to interface with EDI, but the underlying processing would remain EDI. The system could function with traditional methods of EDI and could support additional Internet transport mechanisms such as IP, 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 iden-

---

**Gardner's Handbook of Chemical Synonyms and Trade Names**

11th Edition

**NEWLY REVISED AND UPDATED**

Through ten previous editions, Gardner's Chemical Synonyms and Trade Names has become one of the best known and most widely used sources of information on chemicals in commerce. The Eleventh Edition includes the results of the continuing research underlying this reference work and has seen a major expansion of the information provided for individual chemical compounds. Gardner's provides a wealth of information that can not be found in any other single reference.

**ELEVENTH EDITION INCLUDES:**

- 35,000 Entries
- 5,500 Most Heavily Produced Commercial Chemicals
- 1,500 Chemicals Used in Agriculture
- 85,000 Synonyms
- Over 1,500 Manufacturer and Supplier Addresses

**ENTRIES INCLUDE:**

- Chemical Name
- CAS, EINECS and Merck Numbers
- Molecular Formula
- List of Synonyms
- Uses
- Manufacturers/Suppliers
- Chemical Properties

To order call toll-free: 800-535-9544

E-mail: orders@ashgate.com

Check us out at www.ashgatechem.com
to their contributions. Stated simply, it would be built upon the shared cataloging economic model (participants contribute scanned books instead of bibliographic records), but with the revenues shared by the participants instead of being held by the bibliographic utilities. The price to use the database of collaboratively digitized books would need to be set at a level to encourage use. The profits derived (to be shared by those contributing content) would need to be at a level to encourage the contribution of content. In this sort of model, universities, libraries, and publishers can all be winners.

The strengths of this model includes putting the acrimony and hostility that has characterized our relationship behind us while at the same time increasing the amount of information that one can access without seeking each member’s fiscal destruction. The weakness of this model is that collaboration seems to be an unnatural act. Libraries have seldom been able to cooperate with each other, let alone publishers. Or can we?

So what will our collection development future be like? Three potential models have been reviewed. I am sure each person who reads this could supply others. Just remember, be careful of what you wish for, you might just get it! 🌼

**Chaos**
from page 89

tify 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 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. 🌼

**Back Talk**
from page 94

---

**Your Supplier for European Books**

Please send your orders and inquiries to:

**Erasmus BV**
PO. Box 19140
1000 GC Amsterdam
The Netherlands
Tel.: +31-20-627 69 52
Fax: +31-20-620 67 99
E-mail: erasmus@erasmusbooks.nl
Internet: http://www.erasmusbooks.nl

**Librarie Erasmus**
28, rue Bastou
73011 Paris, France
Tel.: +33-1-44 63 40 20
Fax: +33-1-44 48 14 24

**For books from France**

**ADVERTISERS’ INDEX**

| 25 | ABC Clq | 91 | Bowker | 11 | ISI |
| 43 | ABC Clq | 41 | Cambridge Scientific | 19 | ISI |
| 9 | Academic Book Center | 23 | Carl Corporation | 95 | Kluwer Academic |
| 39 | Academic Press | 65 | Casalini Libri | 47 | Liebert On Line |
| 85 | Accents | 3 | The Charleston Adviser | 81 | MAJORS |
| 2 | Alfred Jaeger | 85 | The Charleston Report | 65 | Martinus Nijhoff |
| 71 | An. Institute of Physics | 69 | Choice | 77 | McFarland |
| 21 | An. Chemical Society | 13 | Dawson Information Quest | 96 | Midwest Library Service |
| 51 | Annual Reviews | 31 | Eastern Book Company | 63 | Portland Press |
| 89 | Ashgate | 33 | Eastern Book Company | 79 | Pivull Libros |
| 63 | ASME | 35 | Eastern Book Company | 67 | Rittenhouse |
| 5 | ATG | 75 | EBSCO Subscription Service | 87 | Schenckso Foreign Books |
| 73 | AUX Amateurs De Livres | 45 | Emery Pratt | 29 | Swets |
| 7 | Baker & Taylor Books | 93 | Erasmus | 17 | UMI |
| 37 | Blackwell’s Book Service | 53 | Greenwood Press | 27 | Wiley |
| 3 | Blackwell's Info Service | 49 | IEEE | 15 | YBP |
| 59 | Book House | 57 | Institute of Physics |  |  |

Ad Manager: Edna Laughrey. Internet: elaquhrey@aol.com
phone: 734-429-1029; fax: 734-429-1711

<http://www.against-the-grain.com> 93