

**"If I could e-speak your language,
I might understand what you e-mean"**

E-Commerce or EDI? What's the Difference Between XML and X12?

by

**Terry H. Floyd, CEO, tfloyd@supgrp.com
tSGi (the Support Group, inc), www.supgrp.com
5010 Doss Rd, Austin, TX 78734
512-266-4400, Fax: 512-266-4401**

"If I could e-speak your language, I might understand what you e-mean"

E-Commerce or EDI? What's the Difference Between XML and X12?

EDI (Electronic Data Interchange), a format for defining documents as items/fields within segments/records, specifies standards, called X12 in North America and EDIFACT in Europe, that allow customers and suppliers, banks, and other institutions, to exchange information. As the Internet continues to replace VANs (Value Added Networks) and proprietary file transfer mechanisms, the need for data exchange between consumers using browsers and legacy systems on the private corporate networks has become more important than EDI was in the 1990's. For readers unfamiliar with the many acronyms used in this paper, an appendix at the end is provided to help in translation and understanding of this already confusing topic.

New ways, like Hewlett Packard's e-speak, to encapsulate the data and specify its format have evolved into ways of using XML (eXtensible Markup Language) for transferring information across the Internet. While some have sought WebEDI, ebXML (e-business), or InternetEDI, others have opted to completely start over. Hewlett Packard's e-speak technology platform relies heavily on standards and open source underpinnings like XML, UML, OBI, W3C, OMG, and CORBA, and relates to more proprietary creations like CBL, eCo, RosettaNet, BizTalk, DCOM, and Java, all of which seem to be broadly supported. Although exactly where EDI standards fit into this murky picture isn't yet clear, the X12 interfaces to the many disparate systems that already exist provide examples of ways to move data between current systems.

Efforts to corner the B2B market

Several of the prominent efforts to replace and/or work with EDI and the Internet through HTML and XML include CBL, RosettaNet, and BizTalk. CBL is the Common Business Library developed by and used at MarketSite.net, CommerceOne's initiative. CBL, like other plans, including classic EDI, correctly assumes that it is easier to connect business systems and services in terms of documents rather than application interfaces which invariably differ greatly. This approach seems to leverage a company's existing EDI usage and allow slowly moving to HTML/XML browser-based tasks.

CommerceNet's eCo (e-Commerce) architecture defines "markets" within networks where businesses provide and use services which conduct transactions that exchange documents containing information items. The lowest layers of this architecture seem very similar to the EDI standards they are supposedly replacing. HP's e-speak fits neatly into this scheme by using XML documents with eCo headers to describe business interaction protocols and mediation.

Another emerging leader is RosettaNet, a set of "industry-wide electronic business interoperability standards." Used initially by the electronic components industry, this initiative has established a common taxonomy for the efficient interfacing of business processes. Their needs-assessment documents correctly point out that the "back-office" interfaces take as much as 50% of the effort when interfacing disparate systems. They see this as a waste of valuable resources that could be used to make new sales or provide new services to their supply chain partners. Using XML, Partner Interface Processes (PIPs) define the framework for more standard interfaces. This approach completely replaces existing EDI processes.

An effort to leverage and preserve much of the existing EDI work of the past twenty years comes from the XML/EDI Group. Their attempts to merge XML with EDI, called XML/EDI, allow for the use of existing document-centric tools in addition to database facilities to manipulate, store, and search transactions in their own e-business framework. It joins the best of both worlds: XML's tagging script attachment, validation, searching techniques, linking, and the use of web authoring tools plus EDI's trading partner profiles, logging and archiving, acknowledgement mechanisms, messaging standards, and application APIs. Sounds too good to be true for companies already using EDI, but this approach may be of little interest to the many smaller companies that have never started their document interchange and B2B interfacing efforts.

Other initiatives include those from Oracle, Simple-EDI/EDIFACT, ebXML, OFX (Open Financial Exchange, OTP (Open Trading Protocol), Health Level Seven (HL7), DataChannel's RIO EIP (Enterprise Information Portal), and the Open Buying Initiative (OBI). There are probably several dozen more projects under way (with more acronyms), each of which promises to talk to or replace all of the others.

Last, but certainly not least, is the Microsoft entry, BizTalk. Microsoft and its partners have, of course, jumped on the XML bandwagon and will attempt to dominate as usual. Microsoft's announcement in June, 2000 that they would make the web their main focus, engulfing everything in the computing world and eventually replacing PC operating systems by making the browser the only end-user platform, will have a huge impact if the Justice Department doesn't divide them up. Even as two companies, Microsoft's applications, like MSWord and Excel, will be available from anywhere in the world via the Internet without installation on every client PC on earth. Since their Internet Explorer is already becoming dominant, giving up the desktop OS space may be moot. By building in XML and other e-business features, Microsoft hopes to make their plug-in features ubiquitous.

So, instead of one or two EDI standards, there are now many different attempts to create new standards and many different communities of users backing methods that they hope will become widely accepted, at least in their own special industry. Deja vu all over again, as Yogi Berra would say. Which of these communities and methods will dominate? Which will survive and which will fall? Only time will tell, but each of their marketing campaigns suggest that if you don't join them now you'll be left behind in the

dust. Will the experience gained in a failed attempt to leapfrog your competition be worth the cost? Maybe? Probably? It is difficult to quantify the value of experience.

Hewlett Packard's e-speak strategy focuses on the World Wide Web Consortium (W3C) community and attempts to extend the work of CommerceNet, RosettaNet, and BizTalk, among others. HP is developing and contributing XML schemas to these communities. Their documentation points out that XML's Document Type Descriptions (DTDs) require significant maintenance efforts, which sounds a lot like the EDI standards effort's requirements.

Years of effort in EDI should not be abandoned

The evolution of electronic business-to-business (B2B) information exchange is a progression from proprietary to open, from specific to general and from more expensive to less expensive. Examples from businesses using applications running on HP3000s show various methods of interfacing disparate systems. The first attempts to transfer data between business systems compared the data in each system, and "mapped" each to the other by creating an intermediate file made up of records and fields common to both systems. As many of these different data exchange systems developed, it became necessary to define different file layouts for the different companies.

EDI "standards" developed by industry-type as the individual searches for non-existent commonality foundered. Apparently, the committees that developed the standards attempted to define every kind of field for every "document" that ever existed. And yet, the disparity between systems was so great that even these bloated designs did not contain enough data types and fields to support each new entrant to the milieu, so the standards continued to expand and a large number of "standards versions" developed, which complicated matters even more. Even so, why ignore all the experience and work that paid off for many large successful companies?

Examples

This author's first encounter with EDI was in 1985 when a MANMAN™ site received specifications for a Purchase Order interface from one of its primary suppliers. The simplest option seemed to be to use the MANMAN™ system's plentiful data elements by modifying the existing Purchase Order printing command to create the outbound file instead of printing a hardcopy PO. This was relatively easy and worked great until the second supplier sent a different file layout. Because of various differences, the Purchase Order file output program had to be modified. That was when we realized that perhaps there would need to be a different version of the program for every future supplier.

The very next year, there was a presentation by Rockwell at the MANMAN™ Conference about standardized EDI for the electronics industry. The speaker touted the importance of "translation software" to ease the task of defining differences between Trading Partners. By the time the 1990 Interex (HPWorld) Conference came around, thirteen companies using MANMAN™ met to discuss EDI. They all agreed that they

wanted to use translation software to turn the EDI "record structure from hell" into a fixed-length ASCII file with various record types that matched the MANMAN™ structures. They hoped that a general-purpose interface program on the HP3000 could then be developed that would work for all of them. Although this idea was a bit naïve, the EDiX/3000™ software product that this author developed as a result of that meeting has been able to handle more than 95% of the needs of inbound Purchase Orders without any changes.

The other 5% of the interfaces required changes to EDiX™ primarily because of specialized ways of handling product prices and scheduled shipment dates. The translator is not able to interface to the MANMAN™ files in order to look up prices or to calculate ship dates based on manufacturing calendars and lead times. That is the job of the interface programs. Sometimes the inbound files contained no prices or dates at all. Numerous methods of handling these fields developed over time, usually with a very time-consuming manual process, because different companies calculate these fields in different ways. Hard coding these sometimes strange practices and formulas into EDiX™ was not a very difficult task once they were understood.

By the time of the 1999 ERP World West Conference, several HP users were talking about their XML interfaces. One who this author interviewed was doing "internal" Purchase Order exchanges with its Singapore business unit using an XML "translator" he had developed himself. He had experience with traditional EDI and used the existing programs as a springboard to the new system development. He assured me that XML was not much different than X12 EDI standards files, but that both needed "translation" and mapping to a flat ASCII file. Having studied XML in detail since then, his conclusion seems correct. Despite the many magazine articles flatly stating that XML will replace the "expensive, hard to program" EDI standards, the interface on each end must still be programmed specifically for the target ERP system. There is no getting around that fact, even though "plug-in" interfaces are said to be available for every major package. There is no plug-in for MANMAN™, and I assume there are many other legacy systems with proprietary methods that will never have a standard plug-in module for XML B2B interfaces for all of the popular transactions.

e-speak from HP

Exactly what is e-speak? Like many other "frameworks", it is difficult to define something that is meant to be all things to all people. The best definition is that it is an e-business platform. It is more than just hardware and software, which allow us to do things (i.e. the "how to do it"). e-speak defines and allows specifying new interactions (the "what to do"). What HTML has done for "eyeballs", e-services does for B2B data exchanges and information flows in both directions. e-speak apparently affects every division of HP and every piece of software they've developed and continue to support.

The heart of the platform is the e-speak Engine that enables agent brokering to electronically establish a relationship with a new trading partner, on the fly. No prior human-to-human negotiations will be required. The plan is for all of their work in

software to be leveraged by adding e-speak functionality and/or interconnectivity, admittedly a lot of development work, so that every product from HP will be able to seamlessly interoperate. e-speak is really all of the new stuff in the middle that performs completely new tasks.

e-speak provides for discovery, negotiation, management, security, and utilization of e-services. After new e-services are developed and deployed the interaction begins. Unattended brokering should involve determining the quality of the service as well as price and availability. This process is called mediation and although there are still some holes in the process, like who can you trust and how do you rate quality of service before you try something yourself, HP is confident that resolution of those issues will come about.

Registry and discovery of services are the basis for some of the functionality provided by the e-speak engines that can cooperate in locating needed resources. In order to interoperate, disparate systems need detailed descriptions of not only the data formats and contents, but also of the processing to truly exchange meaningful information. e-speak provides for those definitions so that once an e-service is located, it can be deployed and exploited. The contract between two businesses must now really define the details of the interface and that is provided by e-speak.

Important products from HP that are not necessarily yet widely accepted but which provide large pieces of the e-speak puzzle are ChangeEngine and OpenView. Also important are the security products like Praesidium and VirtualVault, which add authentication and authorization using digital certificates. HP has innovated in several areas and although standard SSL (Secure Socket Layer) protocols can still be used, their extensions, like SLS (Session Layer Security), even though perhaps confusingly named, provide needed benefits.

The most exciting and tangible evidenced of e-services provided at the Boston e-speak Developer's Conference at the end of June, 2000 was HP's e-dev initiative. Imagine the ability for your development staff to have access to every new release of every new software development tool from *any* where, *any* time on *any* device (*any*³ or *any* cubed). That's what browser access from a single cockpit hooked up to e-dev services will provide. When this is available on a pay-as-you-go basis in a 24 X 7 ubiquitous environment, the power of a computing utility will be realized in a way that will multiply the productivity of software development staffs everywhere. The specific example given by HP was access to new IA-64 platforms for developers wishing to begin re-programming their utilities and applications before the first production machine is ever shipped. By completely outsourcing the programming environment to HP, developers get tools for designing, programming, versioning, project management, profiling, and collaboration as well as system management, capacity/expansion planning and deployment, disk storage, and backups. e-services like these are coming, but not with all the frills in the immediate future. It's going to take some time before all of the popular tools from every software vendor on earth will run in a browser. The e-speak vision is

true, but only time will tell if it becomes a major piece of accepted standards. HP certainly believes it will and has bet the company on it.

Dreaming?

A headline in the July 2000 issue of Knowledge Management magazine boldly states that "XML offers a framework through which businesses can interact despite the inconsistencies among IT platforms, applications, and data formats." Another article, in the June 2000 issue of APICS magazine (aimed at Inventory/Production Control practitioners, not at IT professionals) says "there is no barrier to entry with XML, no sophisticated people needed to implement it, no year of implementation, no high cost implementation... you can put XML connectors or plugs into any software system easily... by comparison, EDI was a totally separate application set."

If you believe these over-simplifications of this complex issue, you will believe anything and need to pay particular attention to problems related below. XML alone is not a panacea and will require lots of work by trained professionals.

Four data conversions instead of one?

The least amount of processing or "fastest" way to transfer data from one system to the other would be to have a single program which reads and selects records or documents from the source system, reformats in memory on-the-fly and writes to the other system: one conversion. The next simplest method would require two conversions -- from the source format to a common format then to the target format. The latter example exemplifies the plans of the original EDI crusaders in the 1970s and 1980s. ANSI X12 and EDIFACT standards define the format of that common intermediate file. In actuality, four programs evolved to handle the EDI process from end-to-end. The first program extracts data from the target system to a flat file. The second program maps and translates the flat file to the target trading partner's selected EDI standard and version. After transfer to the target site, the third program translates the EDI standard file to a flat file appropriate for their application. The fourth program loads the flat file into the target system's database.

The B2B and e-commerce scenario promises a 2-conversion effort, just like EDI did. The straightforward steps are: (1) extract the data from the target database directly into XML, which implies conversion number one. (2) Move the data to the target site. (3) Read the inbound file, converting the data (the second time) from XML and load it into the target database. But what if the existing systems don't support XML or what if both companies already use EDI and have their interfaces working. In that case, four programs are again needed as with EDI. As long as data mapping and conversion are required, why transition from X12 to XML?

And what if the source system doesn't contain all of the data needed by the target system? Or what if the target system's applications don't provide needed functionality to support the business processes for the exchange? An example here is MANMAN's™ lack of data

collection and storage of information about packages, containers, boxes, tares, and pallets for shipments. These fields are required in the Hierarchical Level information for an X12 856 ASN (Advance Ship Notice). Sounds like some heavy-duty modifications or enhancements might be needed immediately. Barriers to data exchange such as these abound and are ignored by EAI and IAI vendor's marketing literature

If these conclusions are wrong and it does soon, and miraculously, becomes possible to easily interface disparate systems, the implications are huge. The debate over best-of-breed software vs. a single-source all-inclusive ERP system will swing again. It will be possible to buy all of the best packages for each individual company's environment and tie them together. That solution would deliver a system custom tailored to the exact needs of a company rather than the compromises of selecting a single system that satisfies some departments and angers others. But don't hold your breath; the reason the pendulum has swung away from best-of-breed integration to single-source ERP is because the integration sounds better in the magazine article than it is in the real world.

Conclusions

Misconceptions about XML replacing X12 and about EDI being replaced by B2B are too common. X12 and XML are just data format standards. EDI and B2B both require customization: to be able to interface with legacy systems and to handle differences between implementations of the same package by different companies in different industries. Both EDI and B2B can use intermediaries like VANs or not use them by going directly between systems over the Internet.

XML will still require the mapping of data from one format to another. But instead of the format definitions being monitored and controlled by a standards body like the ANSI X12 committees that release new but well-documented versions every six months, hundreds or thousands of companies will be maintaining their own DTDs. That means they'll be releasing new version continuously, and unless something unusual happens, versioning will become a nightmare.

The last and most-often discussed problem with X12 as compared to XML is the actual size of the files being transported, processed, and stored. Some erroneous articles have stated that EDI uses fixed-length formats that waste space, but anyone who really knows EDI knows that's not true. In fact, EDI files use *much less* space than XML files because most of the metadata describing the formats is redundantly shipped inside every XML file.

Education and availability of information are the only hopes for influencing the decisions of the many users who will determine which of the proposed standards finally become widely accepted. Just because someone has a better way doesn't mean it will become the standard. Many non-complementary standards have been proposed in the past, such as in the famous BetaMax vs. VHS war, but it is impossible to tell in advance which will win. Users, IT departments, and companies are voting with their dollars, as usual. Even

Microsoft has had to bend in some cases and although HP seems to be squaring off against them again in the internet B2B and e-services area major parts of e-speak will survive in the long run.

Technology moves forward like a spider, moving the leg that furthest behind as the different segments make progress independently. As the vast amount of data exchange experience and knowledge, learned through the use of the more proprietary EDI interfaces, is made available to browser or e-services based interfaces, usage of XML and e-speak will undoubtedly flower. Through e-speak registry and discovery, desired services can be located and negotiations can be carried out. It is the applications-oriented knowledge, business analysis, and project management skills that make interfacing work in the real world. XML and e-speak make it possible to describe these real world situations. If the current trend continues, we'll soon be saying "e-speak what you mean and mean what you e-speak" and the meanings will be "well behaved" because of more precise definitions for object/document vocabularies and services.

Appendix: Definitions and Acronyms

The world of IT, and especially eCommerce, is filled with acronyms, mostly three-letter long, leading to the popular TLA acronym as a joke (see listing below). One listing in this section could be "E-Commerce -- see E-business". Seriously, some more detailed definitions may be needed to progress through this paper, so scan for them on the Internet for the latest meanings.

API	Application Program Interface (business logic)
A2A	Application-to-Application (seamless integration)
B2B	Business-to-Business (see A2A)
CSS	Cascading Style Sheet (output formatting)
DTD	Document Type Definition (document structures),
DOM	Document Object Model (infrastructure)
EAI	Enterprise Application Integration (software glue)
EDI	Electronic Data Interchange (the old/current way to integrate/interface)
ERP/M	Enterprise Resource Planning or Management (big application packages)
HTML	HyperText Markup Language (the basis for web pages)
IAI	Internet Application Integration (see EAI)
P2P	Program-to-Program (see B2B)
RDF	Resource Description Framework (specifications for the Internet)
SVG	Scalable Vector Graphics (graphical output)
SGML	Standard Generalized Markup Language (the grand-daddy of HTML and XML)
TLA	Three-Letter Acronym
VAN	Value Added Network (an pay-by-the-byte electronic mailbox)
UML	Universal Markup Language (systems design)
XML	eXtensible Markup Language (the new lingua franca of the Internet)
XSL	eXtensible Style Sheet (output formatting)