SWP 30/90 ELECTRONIC DATA INTERCHANGE AND ADVANCED INFORMATION PROCESSING - THE WAY AHEAD

ANDY BYTHEWAY
Cranfield School of Management
Cranfield Institute of Technology
Cranfield
Bedford MK43 OAL
UK

(Tel: 0234 - 751122)

and

ROBERT NEELY
Security & Communications
Cranfield Information Technology Institute

ELECTRONIC DATA INTERCHANGE:
THE LONGER TERM EFFECTS ON INTERNATIONAL TRADE

Introduction:

This series of working papers has been prepared as part of the early work in a new programme of research, based at the Cranfield School of Management. The topic for research is the "Longer term effects of Electronic Data Interchange" on business, in the United Kingdom, Europe, and elsewhere in the world.

Contents:

Seven working papers are available, as follows:

Topic	Date
Glossary of EDI terms and acronyms	18th July 1989
EDI: Technical Opportunity or Business necessity?	18th July 1989
Survey of EDI users and service providers in the UK	October 1989
An update report from the "EDI '89" conference in London	6th November 1990
EDI and advanced information processing - the way ahead	15th September 1989
EDI Standards and the Single European Market	11th September 1989
The effects of EDI on the financial sector	18th December 1989

Acknowledgement

We are grateful to the **International Foundation for Artificial Intelligence** and to the **Cranfield Ecotechnology Research Centre**, without whose support and encourgement this project would not have been possible. We are also grateful to the authors who contributed to the different papers; the author and source of each paper is detailed on the cover sheets.

Andy Bytheway,
Information Systems Group,
Cranfield School of Management,
Cranfield Institute of Technology,
Cranfield,
Bedford, MK43 0AL,
United Kingdom.

Telephone: 0234 751122
Telex: 826559 (CITMAN G)
Facsimile: 0234 751806

Introduction

Since 1976 inter-enterprise communications have grown to the point where, today, over six million people use electronic mail and several thousand companies are experimenting with Electronic Data Interchange. U.S. commentators have estimated the access to commercial databases to be a $3 billion business.

Thus, it can be seen that electronic data interchange is becoming a permanent feature of the business environment throughout the developed world. Already innovative organisations are moving beyond the crude and simple versions of EDI which focussed only on creating standard electronic representations for common business documents. As the capabilities of communications and processing technologies become more familiar, there are new demands for wider access to more sources of information, for automation of the access, retrieval and composition functions and for electronic refinement of the raw information using techniques pioneered in defence research establishments and artificial intelligence laboratories.

EDI and Security

Many users, researchers and vendors have pointed out the vulnerability of such EDI transactions and exchanges to electronic monitoring and tampering. A recent market research report looking at trends in EDI in Europe for period from 1989 to 1994 identified data privacy and security as the major worry for EDI users.

This concern is compounded by the government-sponsored drive to Open Systems Interconnection and UNIX-based Open Operating Systems, where security is usually weak or non-existent and many modern threats to confidentiality, integrity and availability have no effective countermeasure in place. As companies migrate to Open Systems, their tactical migration by using gateways to link proprietary networks both extends the area from which attacks can be launched and disables much of the non-standard security sometimes available on proprietary networks.

This has led to security being highlighted as the most important of the three issues that are critical to enabling inter-enterprise communications, the other two being interconnection and the provision of directories.

EDI and Interconnection

Interconnection is needed between different messaging services and products to allow people in different companies to exchange messages without worrying about which type of system they use. Buyers want to compare prices in different electronic catalogues without disconnecting from one system and logging onto another. Customers want to check product availability and order status without concern for which vendor is the middleman. Interconnection is seen as the key to establishing reach and critical mass for electronic messaging systems, although as will be shown, interconnection needs to be complemented with ways to ensure interoperability of the connected applications.

EDI vendors have recognized the need for common user interfaces to permit users access to a large variety of databases using a generic command set. These common user interfaces and common application interfaces are a key feature of the new Open Systems movements such as X/Open, the Open Software Foundation, UNIX International and the Japanese TRON initiative, where the success and rapid take-up of the GUIs (Graphical User Interfaces) such

as the OSF's Motif has been one of the most significant developments. The continuing push to common application interfaces and remote database access standards has been one of the drivers behind the growth of "agent technology" and "information refineries", but the security implications of these technologies has not been given any significant scrutiny.

EDI Directories and Security

As interconnection and interoperability are gradually achieved, the role of messaging directories will become increasingly important. The expectation of users is that these should be comparable to telephone directories in ease of use and vendor-independence. Again, however, there are difficult security issues. Communications users want to protect their internal mailing lists but also want to make their people available for valid business reasons. Already commentators are observing the contradictory nature of users' desires for simultaneous ease of use and tight security. Vendors face the same dilemma. They want the widest possible access to other vendors' directories, but they also want to protect their own customer lists.

Agent Technology

A significant new security threat to financial intermediaries and others comes from the recent advances in AI-based "agent technology". The awareness of the possibilities for security defeating mutants of agents is stemming from the publicity given to the Internet "worm" or "virus" program which penetrated 6000 computers before a solution was found. The chief proponents of agent technology are Apple Computers and Hewlett-Packard but UK and European computer manufacturers are following the same path as they develop AI-controlled "usability sponsors" for their network users.

In broad terms, an agent is an operatorless process that is event-driven and fault-tolerant and that uses artificial intelligence techniques to control other tasks.

The technology has already been applied to automatic search of airline databases for comparative analysis. The need for multiple and simultaneous quotations derived from a number of online databases serving the insurance industry is another area where agent technology will cause a re-appraisal of the broker role.

The key to the further advance of agent technology is the provision of more sophisticated decision making processes and for this capability the providers of agent technology are adopting neural network solutions. This is because neural networks are trained by example rather than programmed by software, which corresponds to the techniques used by the human users of online data services to train others.

EDI and Neural Network Technology

Neural network technology brings completely different concepts to computing. New computing concepts, such as non-algorithmic computing, learning machines, distributed intelligent memory architectures and parallel processing languages based on new mathematical results, require that we leave many traditional computer concepts behind.

Neural computing is a non-algorithmic method of computing which is able to take full advantage of massively parallel computer architectures. The neural networks learn an

application; they are not programmed. Current computer technology depends on software for application development,but the expense of designing, coding and debugging software means that many desired applications cannot be developed, particularly if the anticipated cost of adaptation and maintenance is high. Neural network technology overcomes the software bottleneck through learning since they are trained through examples rather than programmed by software.

The major characteristics of neural networks are:

- o pattern recognition
- o pattern reconstruction
- o adaptability
- o fault tolerance
- o resistance to fuzzy or noisy input

Of particular importance to the agent technology application is the ability of neural networks to be trained through the repeated presentation of examples and to continue learning while the network is being used in an application.

Although the first neural network was proposed in 1957 and early work was carried out at U.S. universities and the University of Tokyo, it is only in the past few years that an explosion of interest has occurred, to the point where the U.S. Defense Advanced Research Projects Agency (DARPA) has recommended that the Department of Defense devote 400 million dollars to neural network research over the next eight years. Much of the impetus to commercial exploitation has come from the availability of hardware accelerators such as those from SAIC, TRW, IBM and HNC. It is significant that this technology is now being fed into the new UNIX-based Open Systems movements by means of submissions in response to the Open Software Foundation's Request for Technology process. This feed-in to de-facto standards making organisations has been actively supported by the Cranfield Information Technology Institute which is the UK's only academic organisation that is a member of the OSF and which is closely involved with the relevant groups in the OSF.

A further boost to the general area of connectionist architectures has come from the commercial success of such machines as the Thinking Machines' Connection Machine and the Advanced Memory Technology Distributed Array Processor. These machines can exploit the power of up to a million processors and are being used for very high speed database searches at speeds of over 4000 million characters per second even from a 4096 processor system. The companies involved have asserted that their technology even at the 4000 processor level will be shrunk onto silicon sufficiently to match PC drop-in card form factors within two or three years.

IT Leverage and Information Overload

Companies are trying to leverage information technologies for radical internal restructuring and competitive positioning internationally.

The modern business realities are:

- the compressed time-to-market phase required for successful products and services;
- the movement toward regional, segmented marketing strategies;
- the increasing incidence of competitive initiatives;
- the impact of new partnering arrangements - often with traditional competitors;
- downsizing, destratification initiatives and structural changes in organisations;

In this environment there is incesssant pressure to deal with more information in a shorter time and to increase the quality of decisions. One solution centres on the new concept of "information refineries". The driver for this development has been the shift in practice to where users' time and attention is now in much shorter supply than the availability of raw information in electronic form.

Information overload is now seen to be a problem for commercial managers where previously it was felt to be only significant in safety-critical systems. Such information overload was cited as the reason for the extent of the Three Mile Island disaster and the erroneous shooting-down of the Iranian jet by the U.S. Aegis system. In business decision-making the consequences of information overload will be poor competitive performance.

Commercial information systems have been likened to third world telephone systems, where the addition of each new phone creates an exponential increase in demand for capacity. In this business environment, "information refining" techniques have been developed that can dramatically reduce the volume of unwanted and unnecessary information.

Information Refineries

Information refining is an electronic, computer-based process that takes undifferentiated volumes of raw information - magazines, newspapers, reports, memos, newsletters, directories and databases - converts them into electronic form, extracts the content units and recombines them into a new form that can be destributed in a variety of ways. The end product of information refining can be sold as a commodity, as a finished good or as an input for another refinery.

The finished product of an information refinery can take many forms: a database, a business letter, marketing report, electronic publication, directory, paper publication, voice message, graphic or even video animation.

Most of the electronic information that today's professional consumes is still in its raw state. Often, it contains only a portion of the information that the user wants, and either he has to search through a series of documents or reports to find out what he wants or he has to query a variety of hard-to-access databases. The information is seldom organised in a fashion that is comprehensive and coherent. In many cases, the cost - in terms of time, effort and intellectual overhead - exceeds the usefulness of the information. As the volume of corporate information continues its exponential growth, it will simply be impractical for users to cull useful data from raw information.

First Generation Refineries:

The first generation of electronic refineries were limited to working with the literal terms of words, keywords and synonym lists. These types of refineries rely on word indexes and the searching and indexing techniques of keyword bases, but suffer from poor selectivity and performance.

Second Generation Refineries:

To move beyond simple keyword filtering-type refineries requires another level of analysis and categorization - typically a mixture of grammatical and semantic categories. In the simplest sense, such systems depend on word categories, word morphological analysis (prefix and suffix cropping), word-sense disambiguation, simple syntactic categories (prepositions, proper nouns, modifiers, modals and so on) and some form of contextual information on the relative position of words.

With this simple, next level of abstraction, it is now practical to build applications that are tailored to look for certain predetermined patterns in the text. Such systems can recognize the equivalence of a variety of ways of stating the same idea but have limited ability for extension. They are best used to convert from one format to another when both formats are highly constrained and scrupulously observed. With the emerging competing EDI formats there may be a role for this type of refinery provided as a dedicated server on a local area network. A side-effect of the use of such refinery servers is to decrease the bandwidth requirement for information flow to the user, who now needs to carry out less visual filtering or local storage and processing.

Third Generation Refineries:

A complete third generation information refinery should be able to take information from diverse sources in diverse formats covering diverse topics and transform them into a basic form that can then be processed into a variety of products and services. Such applications include the population of databases from free text sources, the concrete representation of object models, the generation of composite publications and a variety of alerting and intelligent agent applications. Refineries of this type have been developed primarily for defence and intelligence applications.

As commercial organisations move to fully electronic operational and business environments, a premium will be placed on the selective acquisition and distribution of information. Increasingly, information will not be treated as an overhead item but valued as an input into the work process and refined and recombined to complete a variety of tasks. In the electronic business environment, information refineries will evolve to automate the acquisition adn distillation of information from a variety of external and internal sources.

Migration to EDI

Intra and inter-organisational systems for integrating many of the marketing, logistic and distribution functions of buyers, sellers, manufacturers and suppliers are growing at an enormous rate. They, in turn, are creating enormous volumes of data that have transactional, managerial and analytic value. Not only can companies tie more effectively into their suppliers and customers, but they can use data from their operations to monitor and improve their performance and relationships.

Where EDI has been slow to start, the capabilities of modern high-capacity optical scanning and character recognition devices will enable rapid conversion to electronic representation and the refinery techniques will speed the convergence on standard EDI formats.

A further expansion in EDI's range of coverage is now starting with the appearance of products that incorporate the standards for Office Document Architecture (ODA) and the Standard Generalised Mark-up Language (SGML). Because of the aggressive endorsement of SGML by the U.S. department of Defence, products are already on the market that can scan technical documents and automatically insert the appropriate mark-up commands. As more publishers accept articles and material in SGML form and disseminate their publications in SGML form, there will be a vast improvement in the speed and efficiency of automated classification and retrieval of complex documents.

The maturity of these systems is reflected in the appearance of ODA Conformance Testers such as the TODAC Analyser, which checks for conformance to structure, character content, raster content and geometric content.

Interchange of Design Documents

Related to the use of EDI in the technical area are the advances in EDIF (Electronic Design Interchange Format). This means of communicating design information has already allowed innovative U.S. companies to accelerate their development processes by locating plants around the world in such a way as to achieve continuous working. This was done by siting one plant working on a microprocessor design in California and another plant working on the same design in Israel. Thus at the end of the working day in one plant the design information was sent electronically to the other plant where work was just starting for the day and the process was repeated as the day finished at the recipient plant and their day's reults were communicated back. The success of this technique is now being generalised to take many other factors into account as enterprises are being physically distributed to carefully chosen locations with adequate bandwidth links being planned from the outset.

As strategic alliances are forged between companies with complementary IT strengths, the electronic exchange of design information raises the question of how to adequately limit and protect proprietary processes. One example of this is in the area of alliances between microprocessor manufacturers and test equipment providers. Due to the complexity of modern devices, the behavioural model available in the tester needs to be the same as that used in the manufacturer's company during the design and fabrication processes. This information, however, is key to the vendor's negotiation's with possible second source suppliers and needs to be usable but unreadable by the test equipment provider.

The answer is in the use of cryptographic protection with a system of keys that permit fine control over the dissemination and use of the information. Comparable techniques are now being studied for inclusion in Secure EDI standards and Secure OSI standards.

EDI: New Capabilies and New Risks

EDI is information technology that gives trading organisations new capability. Already in use for national trade in major countries around the world, it gives recognized benefits in cost savings and in faster flows of goods and information together with improved management control.

There are, however, some fundamental business and risks that arise from the potential abuse of EDI.

For example, if EDI is used in a manufacturing environment, customers can modify the order book thus affecting factory loading, stock and re-order levels. Before companies grant access to their own critical operating data, they need to ensure the authenticity of the messages that they receive and react to. Otherwise unscrupulous business competitors or malicious computer criminals could masquerade as legitimate users and wreak havoc with the company's commmercial and management information systems.

Apart from malicious attacks on the system, early EDI trials confirmed the threat from unreliable networking bearer services when some messages were lost completely, other messages were duplicated and some were corrupted. From the point of view of the user the business impact of the distorted information being received was just the same as if criminal actions were involved. Thus we can see that from the users' perspective, the availability, denial of service and quality of service issues should be lumped together, since a loss of availability of a financial information system from poor quality software in the implementation of the network is indistinguishable from a denial of service due to computer hacking or virus insertion.

In the transfer of business-critical financial information using EDI, confidentiality is often of equal importance to integrity. Originators and recipients of messages must be assured that only those people and organisations that are authorised to see the messages should be able to do so. Already senior managers are giving uncertainty over security as their reason for not using electronic mail and office automation.

Apart from these technical issues, there are legal and contractual problems when EDI is used. For example, where a message is originated in one country, held in an electronic database in another country and accessed by the final recipient from a third country, what law applies and at what point in time is the message deemed to have been received from when it may be acted on as a contractual instrument?

At present, most Value Added Network Services rely on distinct usernames and passwords to authenticate the identity of users. Consideration is being given to use of smart card, super-smart card and intelligent token technology but the take-up is slow in the UK and there are disagreements between the French, British and Dutch national body positions on the ISO committee.

For reliability, adaptability and resilience the trend is to rely on implementations of profiles or functional standards that have been agreed by organisations such as the European Computer Manufacturers' Assocation. These profiles are subsetted vertical stacks of Open System Interconnection protocol standards approved by the International Standards Organisation. The obstacle to international take-up of such an EDI base lies in the number of incompatible variations of OSI base layer standards that exist with national networks not being synchronised in their migration and upgrade strategies for revisions of the standards.

The fundamental problem here is that the "open-ness" requirement directly opposes the need for security or as the question has been posed: "how disfunctional does security have to be?".

Secure Messaging and X.400

Because of these concerns and the steady push to X.400 standards as both the international electronic mail standard and the base of EDI, many companies in Europe are developing secure versions of X.400, where the market for X.400 security products has been estimated at up to $2000 million worldwide.

A significant feature of many of these products is their claim of security enhancement expressed in terms of the U.S. defence security classifications taken from the "Orange", "Red" and "Grey" books respectively:

> Trusted Computer System Evaluation Criteria
>
> Trusted Network Interpretation of the Trusted Computer System Evaluation Criteria
>
> Computer Security Subsystem Interpretation of the Trusted Computer System Evaluation Criteria

Despite this many experts are not convinced of their relevance or the suitability of the defence security concepts from these books. This is partly because the concepts and models are old and partly because they are culturally mismatched to the area of commercial computer security.

This has led to the emergence of new national guidelines for commercial computer security such as the UK's "Green Books" and the German "Blue/White" books. The need for European harmonisation of these new standards and the inherent difficulties of certifying end-to-end security in heterogeneous multi-vendor systems has led to new research being funded in this areas both in Europe-wide initiatives and in national initiatives. Government contracts in this area have been awarded to the Cranfield Information Technology Institute where the emphasis is on composability of secure systems and high assurance certification processes for smart card technology.

Although the formal techniques that are being investigated are primarily European in origin, a recent development has been the active and extensive involvement of Japanese experts. In particular, the Hamburg meeting on the 4th July 1989 of the ISO/IEC JTC1/SC21 Working Group 1 saw substantial support and technical contribution from Japan. This work is directly of relevance to the Japanese TRON initiative for Highly Functional Distributed Systems.