NISTIR 5247

Workshop on Security Procedures for the Interchange of Electronic Documents: Selected Papers and Results

Roy G. Saltman
Editor

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Computer Systems Laboratory
Gaithersburg, MD 20899

NIST

NISTIR 5247

Workshop on Security Procedures for the Interchange of Electronic Documents: Selected Papers and Results

Roy G. Saltman
Editor

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Computer Systems Laboratory
Gaithersburg, MD 20899

August 1993

U.S. DEPARTMENT OF COMMERCE
Ronald H. Brown, Secretary

TECHNOLOGY ADMINISTRATION
Mary L. Good, Under Secretary for Technology

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
Arati Prabhakar, Director

Workshop on Security Procedures for the Interchange of Electronic Documents

Table of Contents

Note: Some of the papers included in this volume are copyrighted. The copyright-holders have granted permission to NIST to publish those papers herein but they retain all other rights.

Workshop on Security Procedures for the Interchange of Electronic Documents

Summary and Results

Purpose of Workshop

A Workshop on Security Procedures for the Interchange of Electronic Documents, sponsored by the National Institute of Standards and Technology (NIST) and the Office of Management and Budget (OMB), was held at NIST in Gaithersburg, Maryland, on November 12th and 13th, 1992. The impetus for the workshop and some of the issues that were planned for consideration are discussed in the Announcement paper included in this report.

The fundamental reason for the workshop was that rules for the use of security procedures needed to be devised for the electronic transmission of documents between organizations. This transmission process, usually utilizing electronic data interchange (EDI) standards, is being implemented to reduce paperwork, reduce response times between buyers and sellers, reduce requirements for inventory on-hand, reduce transcription errors, and allow for the computer-based filing and analysis of transmitted documents without the need for re-entry of the data. Applications include purchasing, regulatory and environmental reporting, customs and tariff filings, benefits management, and claims and disbursement information.

The need for security procedures is particularly acute for transmissions in which there are no equivalent paper documents in use. The full benefits of EDI and related techniques cannot be obtained without elimination of the equivalent paper documents, but generally accepted rules for the selection of particular protective techniques for electronic transmissions have not yet been established. Selected security procedures would need to be effective for an environment in which a Federal agency may have a very large number of interchange partners.

Some of the concerns that have been raised about electronic transmission of documents are: (1) whether the true source of a received document could be ascertained, (2) whether there could be confidence that the document has not been altered in transit, (3) whether confidentiality in transit could be assured if required, (4) whether the originator could obtain assurance that the document was received by the intended recipient and, if so, (5) whether it was received prior to a specified deadline. The purpose of security procedures would be to mitigate these concerns.

Requirements for security may be satisfied to varying degrees by differing techniques, some more costly than others. In many cases, the more costly solutions provide greater assurance. In addition,

the Computer Security Act of 1987 requires risk-based solutions for Federal agencies, that is, implemented protective measures must be commensurate with the risk of harm that could result from actual loss of data security or integrity.

A major issue faced by the workshop was the categorization of security objectives. In addition, the workshop needed to consider whether more than one level of risk could be specified in any category. If more than one level could be specified, then different security procedures might be appropriate for the several levels.

Workshop Process

The number of attendees at the workshop was limited, in order to provide a significant opportunity for active participation. The size of the available room, as well as the requirement of active participation, determined the upper limit. Participation was by invitation, and the final count of active attendees included 18 from the Federal Government, two from Federally-funded research and development centers, and 14 from the private sector. One private-sector participant could not attend but was able to submit a paper for consideration. Six observers were also in attendance. Participants were invited to submit papers in advance of the workshop, following their review of the Announcement. These papers were distributed to all attendees, and the papers selected for inclusion in this report are being made available to the public with the permission of their authors.

A reason for requiring attendance by invitation was to assure a distribution of expertise among the needed disciplines and types of experience. Disciplines particularly sought were law, internal control/auditing, security technology, and risk analysis. Current experiences particularly sought were application to EDI within the participant's discipline, or involvement with security and auditability concerns in EDI system implementation or in management of value-added networks. The effort to attract the relevant disciplines and experiences was successful.

Following introductory remarks by Lynn McNulty, Associate Director for Computer Security of NIST's Computer Systems Laboratory, and Bruce McConnell, Chief, Information Policy Branch of OMB, the workshop proceeded through its established agenda under the able moderation of Professor Henry Perritt, Jr. of Villanova University Law School. The agenda provided for a number of presentations by individuals involved with implementation of EDI or related systems. These presentations were made by representatives of the following Federal agencies: U.S. Customs Service (Hugh Davis), Defense Logistics Agency (Capt. Bruce Bennett and William Fox), Environmental Protection Agency (David Schwarz), Securities and Exchange Commission (John Penhollow), and Department of Veterans Affairs (Larry Shows). In addition, discussions of specific issues were presented by panels of two or three participants. These panels

were on the subjects of risk analysis (Bob Jacobson, Paul Ryan and Julie Smith), legal and policy concerns (Michael Baum and Peter Weiss), the role of value-added networks (James Morgan and John Stelzer), auditing of EDI (Paul Moo and Horton Sorkin), and a national public key infrastructure (Charles Chamberlain, David Gill and Jerome Svigals).

Near the end of the first day of the workshop, a set of "prospective propositions" was distributed to the participants. These propositions were a set of statements that, if adopted, could constitute basic security guidance in implementing an EDI system. Included with the propositions was a description of a model situation in which the propositions would be applicable, and a list of security objectives that were addressed by the propositions. The propositions, the model, and the security objectives were discussed by the participants on the second afternoon of the workshop, and recommendations for their revision were made.

The Model

The propositions are to apply in situations in which there are a very large number, i.e., hundreds, thousands, and eventually tens of thousands, of non-Federal organizations or individuals desiring to interchange electronic documents with Federal agencies. The information transmitted between the government and its various "trading partners" may include business, financial, regulatory, administrative, or personal data, having greater or lesser sensitivity depending on the particular application.

To support the interchanges, there are available a number of value-added networks (VANs), each of which has store-and-forward capability and has the function of connecting pairs of end-users for communication sessions. While there are other methods of electronic interchange, the use of VANs is expected to be the most widespread method, and a method that is challenging from a security viewpoint. Each end-user is assumed to be connected to one or more VANs by data communications from a computing system. VAN interconnections make possible transmissions between interchange partners who employ different VANs.

Security Objectives

A list of security objectives is a fundamental precondition in any discussion of the effectiveness of the interchange of electronic documents. Lists of security objectives are provided in the papers included in this report by Michael Baum, Jim Orr, Horton Sorkin, David Schwarz, John Stelzer, and Peter Weiss. The list presented here includes those characteristics upon which general agreement was obtained.

Content integrity - non-alteration within designated portions of a document.

Sequencing integrity - ordering of received documents in the sequence intended by the originator, as repetitions, omissions, and mis-orderings are easily identified.

Confidentiality - prevention of unauthorized disclosure.

Originator authentication - confidence that the purported originator (individual or organization) is the actual originator. There are degrees of confidence.

Recipient authentication - confidence that the document has been received by the recipient designated by the originator. There are degrees of confidence.

Timely delivery - delivery prior to a specified deadline.

Other Security Concerns

Non-repudiation: The workshop was unable to reach agreement as to whether this concept should be additionally included as a separate security objective. Some participants in the workshop noted that the international standard on security architecture, ISO 7498-2-1988(E), in its section 5.2, includes non-repudiation as a security service "which can be provided optionally within the framework of the OSI Reference Model." In the ISO standard, non-repudiation is defined (section 5.2.5) as providing "proof" to one of the parties to the interchange against a false denial of involvement by the other party. The ISO standard, however, contemplates that only cryptographic techniques can provide non-repudiation. Other workshop participants asserted that it is the role of a third-party arbiter, such as a judge, arbitrator or jury, to consider the evidentiary weight of any denial of involvement. In such a situation, the demonstrated comparative trustworthiness of documented audit trails will likely determine which party will prevail. Proposition 14 (see below) reflects this view, and suggests that VANs, under certain circumstances, could provide the necessary evidence for non-repudiation.

System-wide Security: While the propositions presented below primarily concern security of the EDI link between interchange partners, users of EDI must be concerned also with the security of their own and their interchange partner's internal systems. It must be recognized that the EDI process encompasses three broad phases - origination, transmission, and reception. Proposition 6 sets forth a necessary relationship between link security and system-wide security.

Propositions

The security propositions, as revised as a result of the participants' consideration, are a major output of the workshop. While these propositions do not represent a formal consensus of the

workshop, they reflect the sense of the discussions regarding the need for appropriate security in the interchange of electronic documents. Designers and implementers of EDI systems may use these propositions as guidance in establishing security controls for their systems. The propositions, and the papers included in this report, plus detailed system requirements, should point the way to selection of the most cost-effective security technology and internal controls.

Analogy With Protective Techniques For Paper Documents

1. Protective techniques provided for documents in electronic media should be analogous to protective techniques provided for documents of similar sensitivity in paper media.

Commentary to 1: The security objectives identified above are needed for any system of data interchange, even paper. As Peter Weiss notes in his paper included in this report, "... the security protections associated with the traditional use of paper and signatures are so transparent to users and so customary that little thought is given to whether particular transactions require their use." The idea embodied in Proposition 1 is that if data sensitivity remains constant with change to a new transmission system then, with due regard for the technical characteristics of each system, the risk is roughly the same. Thus, if it were not necessary to encypher the data for high assurance of confidentiality in paper, it may not be necessary to encrypt the same data in an electronic transmission. In practice, however, this view is somewhat controversial. Some are of the opinion that "if you are willing to put your information in the mail, you should be willing to entrust it to a VAN." Others are fearful that existing electronic transmission systems are not adequately secure.

Risk and Levels of Security Confidence

2. Confidence that any particular security objective has been achieved may be higher or lower, depending on the protective methodologies and technologies employed. A small number of security levels is envisioned with a variety of techniques available for use to ensure that the specified level of confidence is achieved.

Commentary to 2: A number of papers included in this report propose several security levels. Michael Baum proposes three numbered levels and defines each level with "baseline" characteristics. Peter Weiss proposes four levels, including a non-sensitive level. Jim Orr identifies types of data transmissions in health care applications for each of three risk levels: low, medium, and high. Irvin Chmielewski states that a "balance of high and low tech solutions to achieve EDI security has made the process workable in the real world where large numbers of trading partners are involved." All participants recognized that a "one size fits all"

approach to document interchange security would be inefficient and likely impossible.

3. The level of confidence implemented should be selected according to the risk and magnitude of the harm that could occur if the protection should fail, with due regard for the costs to achieve increased confidence.

Commentary to 3: The wording used here mirrors the risk-based standard embodied in the Computer Security Act of 1987. Roy Saltman, in his paper, compares this concept with that of Article 4A of the Uniform Commercial Code, which calls for "commercially reasonable security." Another concept is that "procedures and technology should be used that are available at a cost lower than the value of the potential risks." Saltman states that these three concepts are consistent in that they recognize that there is an optimal level of expenditures, or "point of diminishing returns" for computer security.

4. Determination of a required level of confidence implies the carrying out of a risk analysis. Part of risk analysis is evaluation of the standards of proof that will be applied in resolving disputes over the integrity and authenticity of information handled by the system.

Commentary to 4: Risk analysis is the subject of the paper by Robert Jacobson. All EDI systems are not the same, Jacobson says, and therefore it is not possible to design a single security program. EDI risks can only be handled effectively with rational risk management. Perfect security is infinitely expensive, but inadequate security leads to unnecessary risk-related losses. A quantitative risk assessment needs to be performed because the cost of security measures is stated in monetary terms, Jacobson states, and installing a security measure is not prudent unless its benefit outweighs its cost. With regard to resolving disputes, Jacobson proposes that quantitative risk assessment techniques could be used to analyze the cost/benefit of a service to automatically log messages between interchange partners; the log could serve as a neutral audit trail. Participants pointed out, however, that risk is often difficult to quantify and that a qualitative risk analysis may be adequate, albeit somewhat more subjective. It was suggested that adequate practical experience with various types of interchanges will soon be available to enable risks to be assessed more confidently.

5. A particular document type may require higher levels of confidence for some security objectives than for others. Additionally, the required level of confidence may vary, for the same document type, according to the risk inherent in the data being interchanged, e.g., a high-value purchase order versus a low-value purchase order.

Commentary to 5: Jim Orr, in his paper, specifically identifies four separate security categories, and three levels of risk in each category. In this typology, the transmission of certain types of data could have high risk in one category and low risk in another. Both Orr and Peter Weiss point out in their papers that the monetary value with which the interchange is concerned has an effect on risk, with low value interchanges less risky than high-value interchanges. This concept already exists in the area of paper-based transactions. Specifically, the Federal Government, in its Federal Acquisition Regulation, identifies a procurement for less than $25,000 as a "small purchase" and requires less documentation and simpler procedures, often including solicitation of prices telephonically, for completion of the transaction.

Total System Security vs. Interchange Link Security:

6. The implemented security of a data interchange system should be consistent with the security of the originator's and recipient's internal systems. Good access control to the internal systems will add to confidence in the authenticity of transmitted documents, and effective controlled access to sensitive databases will add to confidentiality when that is required.

Commentary to 6: Irvin Chmielewski, in his paper, stresses "application systems that will utilize EDI" rather than just EDI. This leads him to consider overall security, not just data link security. He states, "... applications, both client/server and host-based, at the manufacturers and their trading partners, must provide full security of all information and resources." In the paper by Horton Sorkin, the security of the data interchange is related to the internal control structure as seen by an auditor of a client's system. Sorkin states that the standard audit objectives for each transaction do not change with EDI; the same concerns are translated to the new technology. "The loss of controlled and prenumbered paper stock is supplanted by concerns over counters, date-time mechanisms, and authorization and internal access controls. It does not appear to make much sense to be concerned about EDI transmission security if internal controls at the business application level are inadequate," Sorkin states.

Non-cryptographic Internal Control and Security Techniques:

7. There are non-cryptographic techniques that can be used to provide confidence in document integrity and originator authentication at a lower level than could be assured with cryptographic techniques. There are no non-cryptographic techniques that can provide high confidence of confidentiality in document interchange.

Commentary to 7: Peter Weiss provides a list of computer security techniques applicable to EDI, presented in generally ascending order of security strength. The lower-level techniques presented are non-cryptographic. Weiss points out that the list of

techniques is generally consistent with that contained in the American Bar Association's Model Payments Agreement and Commentary ("Model Agreement"). 32 Jurimetrics 601 (Summer, 1992). Similarly, Michael Baum's Security Baseline - Level 1 includes only non-cryptographic techniques.

The issue of confidentiality is the most difficult analytically. On the one hand, the business community has accepted the relatively low level of confidentiality provided by the postal system as adequate and the risk of their information being improperly divulged as adequately low. Therefore, it must be demonstrated that the risk of such disclosure on electronic networks is also acceptably low. On the other hand, only cryptographic techniques can provide high confidence of confidentiality. While the installation costs of encryption may be low, the maintenance costs (especially at the administrative level) is an impediment to the use of this technology. Moreover, data encryption is not now in wide use for commercial transactions, even those that are considered sensitive.

8. Some form of standard receipt acknowledgment should be returned to the originator within an agreed-upon time period when the document is to further an implied or explicit contractual or legal relationship. Such acknowledgment confirms receipt.

Commentary to 8: Horton Sorkin states that "The sending of a [functional acknowledgment] without errors reported, or the sending of an Application Advice ... or a specific application response, is an implicit acknowledgement that the security process was successful." Functional acknowledgement is included also by Peter Weiss in his list of security techniques. Michael Baum cites the Model Agreement as stating that "the receipt by the sender of an acknowledgment from the recipient shall constitute conclusive evidence that the subject communication was received and is syntactically correct."

9. An additional method of lower-level authentication is the inclusion in the document of passwords or codes known only to the interchange partners.

Commentary to 9: The use of imbedded references is included by Peter Weiss in his list of security techniques. This technique can be conceived as an extension of a log-on process. The use of references that validate the message originator is similar to a second log-on being done on the interchange partner's system, subsequent to an initial log-on to the originator's system.

10. In cases of lower risk, confidence in message integrity may be obtained by reasonableness checks on data values, and by the matching recalculation by the recipient of real and hash totals that cover the essential parameters of the document. Additionally, document integrity may be further assured by the successful retransmission of its essential content back to the originator.

Commentary to 10: In a commercial environment, data integrity may be much more significant than confidentiality. Reasonableness checks and recalculations on data values make sense, even if there is no concern whatsoever about the transmission link. Data values may be in error due to mistakes at the interchange partner's computer, and the risk of monetary loss in making decisions on incorrect data may be high. Weiss reports that the Model Agreement states that "Consistency checking of the payment amount with prior transactions or customer profiles" is a verification technique. Retransmission back to the originator may be simple to execute. It may only require the turnaround of the message as received, with the addition of an indication of acknowledgment meaning, for example, "We have received your message stating ..." or "We agree to carry out the request in your message that ..."

Cryptographic Techniques for Confidence Under High Risk:

11. In cases of highest risk, the use of cryptographic techniques is necessary to assure document integrity and originator authentication. Techniques using public key encryption, i.e., digital signatures, should be considered when the risk for loss of integrity or failure of authentication exceeds the cost associated with the use of such techniques.

Commentary to 11: All of the authors who listed types and strengths of techniques specified cryptographic techniques as the strongest and most applicable to messages of the highest risks. However, a method of general key management and distribution at a reasonable overhead cost for the large numbers of expected interchange partners would be necessary to a make cryptographic techniques an integral part of electronic commerce.

Originator Accountability - Organization or Individual:

12. In assessing risk and developing security plans, agencies should be explicit about whether legal responsibility is imposed on the organizations or the individuals serving as interchange partners; accountability should be sought at the appropriate level.

Commentary to 12: David Schwarz, in his paper included in this report, makes an important distinction between organizational and individual accountability. This distinction may be necessary, says Schwarz, if criminal enforcement action is to be taken against individuals. It could be essential in such cases to have strong evidence that the specific individuals believed to be responsible for criminal conduct be positively identified as being the source of incriminating electronic documents. Conceivably, for this type of prosecution to be successful, claims that an electronic signature was 'forged' might have to be refuted.

13. Costs to assure originator authentication are likely to be less if accountability can be associated with organizations rather than with individuals.

Commentary to 13: Accountability to an individual at the originating location may require technological and administrative techniques of a more complex and costly nature than accountability to the organization. In addition, at the receiving location, the need for individual accountability of a received message implies a detailed concern for proper authorization, again a more complex requirement. Organizational accountability should be adequate for the majority of commercial and administrative interchanges which do not have direct regulatory or enforcement implications.

Use of VANs as Trusted Third Parties:

14. If VANs could be established to act as neutral and trusted third parties, they could be employed to provide a high level of confidence in originator and recipient authentication when accountability is determined to be with the organizations interchanging data. When trusted VANs are used, placement of a message in the recipient's mailbox by a VAN, without any disclaimer, constitutes assurance that the message comes from the purported sending organization; a report of this receipt back to the originator similarly constitutes assurance that the message was received by the specified recipient organization.

Commentary to 14: VANs transmit messages that they obtain at senders' mailboxes and deliver to recipients' mailboxes. They can provide complete audit trail information from mailbox to mailbox if needed VAN interconnections are effectively implemented. With adequate user access-control to mailboxes, this process may provide sufficient authentication and timely delivery information for many, if not most, commercial and administrative applications. John Stelzer, in his paper, discusses the satisfaction of security objectives by VANs. He says, "Judicious and consistent use of the broad array of security and control tools available in the EDI standards and from EDI networks can provide a high level of reasonable assurance that all six of the security and control objectives are being met in all but the most critical transactions."

Proposals for NIST Activities

In the course of their deliberations, the participants undertook a discussion on the topic of future activities for NIST. Some of the proposals that were made are listed here, without valuation of their merit. In addition, Jim Orr has made a number of recommendations in his paper as to what the Federal Government could do on subjects related to EDI in health care.

Proposal: NIST should evaluate risk-analysis methodologies related to electronic commerce issues, with a view towards developing more detailed guidance for agencies that do risk analysis.

Proposal: NIST should undertake a careful investigation of the concept of trusted intermediaries and the mechanism of how they come to be trusted. Users will be reliant on the intermediary's registration and certificate management activities. This may put an enormous burden on some kind of infrastructure that does not exist right now. Models that could be used include common carrier regulation, licensing activities, and the notary public function.

Proposal: NIST should encourage the directions in which development of commercial products for security should be headed. This would involve defining the best possible security environment in terms of products, security features, and cost, obtainable within the near-term and mid-term.

NIST will consider these proposals.

UNITED STATES DEPARTMENT OF COMMERCE
National Institute of Standards and Technology
Gaithersburg, Maryland 20899
Computer Systems Laboratory
March 24, 1992

Announcement of a Workshop on

SECURITY PROCEDURES FOR THE INTERCHANGE OF ELECTRONIC DOCUMENTS

November 12 and 13, 1992

1. Immediate Origin of the Workshop

A need for accelerated development of Government-wide security procedures for the interchange of sensitive-but-unclassified electronic documents has been recognized by the National Institute of Standards and Technology (NIST) and by the President's Office of Management and Budget (OMB). A recent General Accounting Office (GAO) Decision (B-245714, 12/13/91), directed to NIST on the subject of the use of electronic data interchange (EDI) technology to create valid obligations, suggests that specific security procedures are desired that would protect interchanges as a function of risk. Risk-based implementation of protective techniques is a concept set out in the Computer Security Act of 1987. The GAO Decision, an advisory opinion, did not (nor would it be expected to) identify such specific procedures.

Specific "generally accepted" procedures (as in "GAAP", generally accepted accounting procedures) have not yet been developed by the appropriate security, audit, accounting, and legal professional organizations, because of the newness of the application. These organizations are, for example, the Information Systems Security Association, Institute of Internal Auditors, Electronic Data Processing (EDP) Auditors Association, American Institute of Certified Public Accountants, American Bar Association, etc. Consequently, there is no authoritative basis for immediately promulgating acceptable Federal procedures.

It was agreed that the development of procedures should begin with a workshop on the subject. Those persons who would be asked to participate would be those most knowledgeable and experienced in the field, and they would be selected so as to represent the various pertinent professional competencies, and well as both the public and private sectors. The function of the workshop would be to develop a consensus on the types of protective techniques that should be implemented as a function of risk. The output of the workshop would provide a foundation for the development of a Federal implementation requirement, possibly a Federal Information Processing Standard (FIPS).

2. Issue Background

The meetings at OMB of the interagency Electronic Signature and Message Authentication Task Force over the past two years have highlighted the issue of protective techniques needed in the interchange of electronic documents. In this context, "electronic document" means a predefined and structured message, transmitted as part of a business activity, that is interchanged electronically between computers of different organizations. The predefined structure makes possible the use of computer software for the composition and decomposition of the messages, thereby eliminating much human intervention in message processing. Before the implementation of electronic interchange, such a document would have been prepared in paper and transmitted manually, usually by mail. The new process is generally referred to as EDI, and to be consistent with FIPS Publication 161 on EDI and to make maximum use of available software and system designs, the messages should be composed using the suite of standard formats developed by the national or international EDI standards bodies.

2.1 Benefits of EDI Over Paper-based Interchange

The primary thrust of EDI has been in private-sector commerce, for use in procurement of goods by manufacturers from material and component suppliers and by retailers or wholesalers from manufacturers or other suppliers. Significant use is made of EDI to replace paper media for specific documents such as requests for quotes (RFQs), purchase orders, shipping and receiving notices, bills of lading, invoices, and remittance advices. A transmission of actual monetary value using electronic funds transfer (EFT) may be combined with an EDI remittance advice (which is just data, as are all EDI messages) to pay a supplier through its bank and inform the supplier of the payment, in a composite message.

EDI has gained wide acceptance in the private sector through cost reduction, elimination of paperwork, reduction of the transcription process and consequent transcription errors, reduction of response time in notifying trading partners of the need for re-supply, and better management of inventory (using so-called "just-in-time" control) that the speed-up makes possible.

2.2 Applications of EDI in the Federal Government

Interchanges of electronic documents by Federal agencies will be, in the near term, primarily with private sector organizations, but eventually more widely with the public at-large. While the most widespread Federal application at this time is for the procurement process, as in the private sector, other Federal applications are Government-specific. These additional applications will concern the submission of tax information, and of many types of documents that certify that specifically identified activities now meet or will meet the requirements of Federal law or regulation. In gener-

al, a document is appropriate for electronic replacement when it can be simply structured in a standard format amenable to computer processing, the volume of interchanges is high enough for long-term fiscal savings, and a substantial portion of the Government's interchange partners for the document have access to the necessary technologies.

2.3 The Security Issue in Private-Sector Acceptance of EDI With a Government Partner

Widespread acceptance of EDI will only occur if it has, and is perceived by the business community to have, at least the same level of security as the existing paper-based system, and if the administrative overhead and additional costs, if any, in achieving the same or a higher level of security can be justified by the benefits obtained. To the extent that EDI uses telecommunications or magnetic media to transmit the data, the implementation of security must focus on the confidence that users will have that the new methods meet business requirements; specifically, that "reasonable" assurance is achieved that document authenticity, confidentiality, and integrity are preserved to the extent needed. The concept of "commercially reasonable security" is embodied in accepted commercial codes.

Electronic and paper media share many of the same security risks. However, the security characteristics associated with the traditional use of paper and signatures are so transparent to users and are so routine that little thought is given to them. In paper media, confidentiality is ensured by placing the message in a secure container - an envelope, perhaps enhanced with a seal - which is then delivered to its destination by a trusted courier. The authenticity of the document is likewise ensured by physical means - handwritten signatures, seals, notarizations, etc. Statutory and regulatory requirements routinely specify that communications be "in writing," "signed," "verified," or "acknowledged." These have become so ubiquitous that most routine paper-based communications, particularly forms, customarily contain a requirement for a signature even in the absence of any specific legal or administrative directive that an original autograph signature actually be affixed.

The electronic replacements for these existing security features, including their administrative overhead and cost implications for Government's private-sector interchange partners, are the subject of this workshop.

3. Examples of Government Interchanges

The following examples are intended to provide an indication of the volume of particular interchanges, the types and characteristics of interchanges, and the types and numbers of partners that will be

experienced by Federal agencies as electronic interchange becomes prevalent.

The Defense Department (DoD) estimates that its universe of potential vendors for procurement via EDI includes about 300,000 companies, most of which are extremely small. The overwhelming majority (over 98%) of contract actions by DoD are for orders less than $25,000, which are called "small purchases" in the language of the Federal Acquisition Regulations (FAR). These purchases require less formality for approval. Typical small purchases would include such requisitions as a carton of salad oil containers for a military commissary, or a gross of tongue depressors for a military field first-aid station. Of the 700,000 orders placed by DoD under General Services Administration multiple-award schedules in fiscal year 1988, 667,000 were for small purchases.

The Environmental Protection Agency (EPA) is currently implementing two applications of X12 transaction sets (i.e., national standard message types), called Hazardous Waste Manifest, and Report of Hazardous Waste Activity, in cooperation with 40 states. The states expect to receive a total of about 4,000 manifests a week from a universe of about 100,000 organizations. The EPA estimates that for all its programs, there might be as many as 1 million organizations that might be submitting some type of form, from time to time.

The Health Care Financing Administration (HCFA) is automating the receipt of financial information from Health Maintenance Organizations (HMOs). The information is financially sensitive. HCFA expects to receive about 1900 such reports per year, but each form is currently about 25 to 30 pages of data, on both sides of the page.

The Federal Communications Commission receives about 300,000 applications per year for a certain class of mobile radio transmitters typically used at construction sites and other locations where persons at a distance from each other must communicate in order to work together. Applications, typically received from a universe of about 100,000 organizations, are amenable to use of EDI.

The Internal Revenue Service is developing a program that includes the electronic filing of tax and information returns. Clearly, the number of organizations and individuals eventually participating could be on the order of millions, and the information received would be financially sensitive.

4. Workshop Considerations

4.1 Summary of Assurance Requirements

In many cases, issuance of a document by the originator and its transmittal to the recipient constitutes establishment of an obligation or certification. Assurance of originator authentication may be required, and a non-repudiation capability, although probably not needed for a majority of document types, is nevertheless a potential requirement. Additionally, assurance is often needed that a document has not been changed in transit, and has actually reached its recipient in an appropriate time interval or in advance of a particular time. A further concern is with confidentiality. Some documents contain company-proprietary data, such as official bids as responses to RFQs, and information about company financial conditions; other documents may contain personal data, such as medical information. A summary of available EDI security techniques is contained in the NIST Computer Systems Laboratory (CSL) Bulletin on Security Issues in the Use of Electronic Data Interchange, June, 1991.

4.2 Risk Categories and Risk Levels

A possible workshop outcome is that several risk categories, for example, authentication, integrity, confidentiality, and time-sensitive assurance of delivery, will be identified, and that it will be agreed that each document type should be rated separately for risk in each category. In addition, several risk levels could be defined, thereby making possible the specification of varying strengths of protection as a function of level.

The specification of risk categories and levels, if eventually adopted, would not extend to the evaluation of particular electronic documents. It will remain the responsibility of each Federal agency to specify the particular level and category of each electronic document that it interchanges. Thus, questionnaires and delivery orders could be determined by the agencies involved to have a low level of authentication risk, while tax information submitted by a filer could be established to have a high level of risk in that category, as well as a significant confidentiality risk. A particular type of electronic document may have different levels of risk for different categories or applications. For example, purchase orders used for low-dollar-value transactions may be determined to have low authentication risk, while otherwise identical purchase orders used for high-value transactions may carry a greater risk.

4.3 Criterion of the Computer Security Act of 1987

For establishment of particular protective techniques, the likelihood of occurrence of the several types of exposures must be considered and it must be determined, in a general sense, what expense

is reasonable for protection against them. This is the essential requirement of the Computer Security Act of 1987, which assigns to NIST responsibility for developing standards and guidelines needed to assure the cost-effective security and privacy of sensitive information in Federal computer systems. The Act states that implementation of protective techniques shall be "commensurate with the risk and magnitude of the harm resulting from the loss, misuse, or unauthorized access to or modification of the information ..." For messages where confidentiality is a major issue, it might be determined that it is worth the expenditure to protect against eavesdropping if that can be done at a reasonable cost and with reasonable administrative simplicity. For messages for which assurance of delivery is an issue, or for which undesired changes in numerical values contained within the message would be seriously detrimental, appropriate protections may be recommended, again using reasonable cost and administrative simplicity as criteria.

4.4 Selection of Specific Protective Techniques

The kinds of protective techniques associated with documents embodied in paper media must be appropriate to the media, and analogous electronic techniques at the same level of protection must be appropriate to the new medium. In electronic media, alteration of transmissions as a function of electrical noise or equipment malfunction is a possibility that does not arise with paper media. In addition, passive eavesdropping is a possibility, equivalent to someone steaming open an envelope to read the contents in paper media, but possibly more easily done without detection. Active (and malicious) modification of messages also has some possibility of occurrence.

Misdirection of messages, longer-than-expected delay, or the total loss of a message in the delivery system are also possible risks to be considered; these arise with manual delivery as well as with an electronic system. However, due to the computer-based nature of an electronic system in which there is storage as well as the ability to establish and maintain audit trails, tracing and recovery in a short period of time are much more likely than with a paper-based delivery system. Selected protective techniques should be able to take advantage of these system characteristics.

There is the possibility of using techniques of both types of media to satisfy requirements, and this is being done now. Trading partner agreements may be executed in paper, with manual signatures, in which the agreements provide for acceptance of each partner's electronic transmissions without regard to the presence of a signature on the electronic messages. Whether this is feasible with the very large number and, in many cases, anonymity of the Government's interchange partners is a question for the workshop.

An essential question for this workshop is: what are appropriate electronic protective techniques for the several risk categories

and levels, given the transmission environment, and how should they be implemented? It is hoped that a consensus will be reached.

As an example of a potential recommendation, use of hash totals to assure the integrity of numerical values could be proposed by the workshop to be appropriate for a low level of risk, while use of cryptographic techniques could be recommended for a high level of risk in the same category. Similarly, it could be proposed that for a high level of risk of assurance of correct delivery, a full-content acknowledgement should be sent while for lower levels, either a functional acknowledgement or none at all would be appropriate.

Workshop participants must keep in mind the fact that they are considering protective techniques for only part of a total system: the internal systems of the various originators and recipients of electronic documents are beyond the scope of this effort. Nevertheless, questions of comparative security levels may be considered, as the levels of protection achieved in internal systems, particularly for authentication and confidentiality, may affect the requirements for levels of protection in interchange of electronic documents. A possible recommendation of the workshop could concern the relative protection levels of internal and interchange subsystems for any or all of the risk categories.

4.5 Value-Added Networks (VANs)

The use of a private company serving as a public carrier is likely to be more widespread for general electronic delivery than is now current for documents in paper media. From the point of view of the necessary protective techniques, the capability of the carrier to carry out a function of trust, as well as liabilities for failure to do so, are important considerations. The ability of the carrier, serving as a VAN, to provide audit trails of time and place of origination and receipt is fundamental, especially if more than one VAN must be involved to finally deliver the message. It is hoped that the workshop can provide guidance in the security, audit, and legal arenas as to what capabilities should be requested of such private carriers, in terms of their conduct towards the messages carried, their responsibilities to their customers, and the information about the messages carried that must be made available to the originators and recipients.

4.6 Workshop Format and Results

Workshop participants will be drawn from experts in various pertinent areas: EDI system implementation, computer security, EDP auditing, law of electronic commerce, VAN systems management, etc., and from both the public and private sectors. Each participant will be asked to prepare a short but succinct paper on a pertinent topic and to present that paper to the other participants. Papers on experience with EDI in the private sector and its translation to

Government interchange, use of VANs as trusted third-parties, and applicability and acceptability of specific types of security techniques, will be particularly needed.

Papers on private-sector experience should pay special attention to achievement and assurance of trust and acceptability of received messages among trading partners, and acceptance by internal and external auditors, and legal counsel, of system controls and system validity. Papers on applicability of specific types of security techniques should pay special attention to the issues of administrative complexity, security benefits, and costs per subscriber of implementing those techniques in situations in which Federal agencies receive messages from hundreds of thousands, possibly millions, of partners with limited resources. Improvements in cost-effectiveness of protective techniques over time, due to advanced technical developments and to economies of scale in manufacture and use, must be factored into the selection process.

It is expected that the workshop, to be held at NIST, will initially convene in a plenary session, and presentations of interest to everyone will then be given. Special groups may then be formed around particular issues such as definition of risk categories and risk levels, use of VANs as trusted third-parties, and specific security techniques. Groups may make tentative recommendations, and then bring them back to another plenary, where the final workshop output will be generated. The presented papers, a summary of the discussions, and specific workshop recommendations will be published by NIST.

5. Participation in the Workshop

Expressions of interest in participation in the workshop are solicited, and should be communicated to NIST by August 1, 1992. Invitations to participate will be extended shortly thereafter. Questions about the workshop should be directed to:

Peter Weiss
Office of Management and Budget
phone: 202-395-4814

or

Roy G. Saltman
NIST
B154 Technology Building
Gaithersburg, MD 20899
phone: 301-975-3376
E-mail: saltman@ecf.ncsl.nist.gov

WORKSHOP ON
SECURITY PROCEDURES FOR THE INTERCHANGE OF ELECTRONIC DOCUMENTS
LIST OF PARTICIPANTS AND OBSERVERS

[Information correct as of Nov. 12, 1992]

Jack Bartley (1)
Director of EC
Dept. of Defense CALS/EC/EDI
5109 Leesburg Pike - Suite 701
Falls Church, VA 22041
Ph: 703-756-8471

Michael Baum (2)
Principal, Independent Monitoring
33 Tremont Street
Cambridge, MA 02139
Ph: 617-661-1234

Capt. Bruce Bennett (3)
U.S. Navy Program Manager, EC/EDI
Defense Logistics Agency
Cameron Station
Alexandria, VA 22304
Ph: 703-274-6031

Dennis Branstad (4)
Senior Research Fellow - Computer Systems Laboratory
A216 Technology Building
NIST
Gaithersburg, MD 20899
Ph: 301-975-2913

Michael Buckler (5) Observer
Special Services Officer
Federal Retirement Thrift Investment Board
805 15th Street, NW - Suite 500
Washington, DC 20005
Ph: 202-523-8028

Robert Campbell, CEO (6)
Advanced Information Management
12940 Harbor Drive
Woodbridge, VA 22192
Ph: 703-643-1002

Charles Chamberlain (7)
General Manager, Corporate Applications Division
US Postal Service
475 L'Enfant Plaza, SW
Washington, DC 20260
Ph: 202-268-5262

Irv Chmielewski (8)
Business Planning Specialist
EDS
Mail Stop 6A
650 Tower Drive
Troy, MI 48007
Ph: 313-265-9258

Hugh Davis (9)
Director, Security and Standards Division
US Customs Service
1301 Constitution Avenue, NW - B146
Washington, DC 20229
Ph: 202-927-0185

William Fox (10)
Director, Information Systems Facilities
DLA Systems Automation Center
Attn: DSAC-F
PO Box 1605
Columbus, OH 43216-5002
Ph: 614-692-8200

Harold Frohman (11) (Observer)
Research Fellow
Logistics Management Institute
6400 Goldsboro Road
Bethesda, MD 20817-5886
Ph: 301-320-7286

David Gill (12)
Group Leader
MITRE Corp.
7525 Colshire Dr.
McLean, VA 22102
Ph: 703-883-5926

Tom Hausken (13) (Observer)
Analyst
Office of Technology Assessment
U.S. Congress
Washington, DC 20510-8025
Ph: 202-228-6783

Ken Hoffman (14)
Director of Information Resources Policy
and Standards
Department of Veterans Affairs
Mail Stop 72; Room 244
810 Vermont Avenue, NW
Washington, DC 20420
Ph: 202-233-5434

Paul Hoshall (15)
Director, ADP/IRM Audit Division
Office of Inspector General - Mail Stop 52D
Department of Veterans Affairs
810 Vermont Avenue, NW
Washington, DC 20420
Ph: 202-233-5637

Robert Jacobson, President (16)
International Security Technology, Inc.
99 Park Avenue - 11th Floor - US Re
New York, NY 10016
Ph: 212-557-0900

Thomas C. Jones (17)
Director of Engineering
Lemcon Systems, Inc.
2104 West Peoria Avenue
Phoenix, AZ 85029
Ph: 602-944-1543

John Penhollow, Director (18)
Office of EDGAR Mgt.
Securities and Exchange Commission
450 5th Street, NW
Washington, DC 20549
Ph: 202-272-3900 x3001

Jerry Malitz (19) (Observer)
Statistician
Nat'l Center for Educational Statistics
555 New Jersey Avenue, NW
Washington, DC 20208-5651
Ph: 202-219-1364

Christopher Martin (20)
Assistant Director, AFMD
General Accounting Office - Room 6009
441 G Street NW
Washington, DC 20548
Ph: 202-275-9481

F. Lynn McNulty (21)
Associate Director, Computer Security
NIST
B154 Technology Building
Gaithersburg, MD 20899
Ph: 301-975-3241

Brent Melson (22)
Information Systems Auditor
NASA - Code W
Washington, DC 20546
Ph: 202-453-2943

Paul Moo (23)
Senior Manager
Price Waterhouse & Co.
15301 Dallas Parkway - Suite 300
Dallas, TX 75248
Ph: 214-386-9922

James Morgan (24)
Manager, GEIS Security
General Electric Information Systems
401 N. Washington Street
Rockville, MD 20850
Ph: 301-340-4906

Glen Mules (25)
Industry Consultant
Stratus Computer
17 State Street
New York, NY 10004
Ph: 212-530-0662

Charles Noble (26)
Electronic Information Security Consultant
Digital Equipment Corp.
111 Powdermilk Road - MSO 2-2/A3
Maynard, MA 01754-1499
Ph: 508-493-8728

James Orr (27)
Director, Support Services
Blue Cross of California
21555 Oxnard Street - Mail Station G-K
Woodland Hills, CA 91367
Ph: 818-712-6764

Prof. Henry H. Perritt, Jr. (28)
Professor of Law
Villanova University School of Law
Villanova, PA 19085
Ph: 215-645-7078

Paul Ryan (29)
Analyst
Internal Revenue Service
Electronic Management System Project Office (ISD:D:EMS)
Washington, DC 20224
Ph: 703-235-0242

Roy Saltman (30)
Computer Scientist - Computer Systems Laboratory
NIST
B154 Technology Building
Gaithersburg, MD 20899
Ph: 301-975-3376

David Schwarz (31)
Chief, Information Policy Branch
Environmental Protection Agency
401 M Street SW - Room 3313
Washington, DC 20460
Ph: 202-260-2706

Lawrence Shomo (32) (Observer)
Manager, Automated Information Systems
NASA - Code BAB
Washington, DC 20546
Ph: 202-358-1041

Larry Shows (33)
Acquisition EDI Coordinator
DVA Data Processing Center
Department of Veterans Affairs
1615 East Woodward Street
Austin, TX 78772
Ph: 512-326-6019

Miles Smid (34)
Group Manager - Computer Security Division
NIST
A216 Technology Building
Gaithersburg, MD 20899
Ph: 301-975-2938

Julie A. Smith (35)
Research Fellow
Logistics Management Institute
6400 Goldsboro Road
Bethesda, MD 20817-5886
Ph: 301-320-7369

Horton Lee Sorkin, Ph.D. (36)
Professor, Howard University School of Business
3806 North Nelson Street
Arlington, VA 22207
Ph: 703-243-1652

John L. Stelzer (37)
Senior EDI Consultant
Sterling Software - ORDERNET Services
PO Box 7160, 4600 Lakehurst Court
Dublin, OH 43017-0760
Ph: 614-793-7046

Jerome Svigals (38)
President, Jerome Svigals, Inc.
221 Yarborough Lane
Redwood City, CA 94061
Ph: 415-365-5920

Geoffrey Turner (39)
Technical Director
Mantech Strategies Assoc., Ltd
445 Baldrock Road
Kalispell, Montana 59901
Ph: 406-756-9202

Gilles Vezina (40) (Observer)
Office of Comptroller-General
Government of Canada
300 Laurier Avenue West - 8th Floor
Ottawa, Ontario K1A-1E4
Canada
Ph: 613-957-7047

Peter Weiss (41)
Senior Policy Analyst and Attorney
Office of Management and Budget
Executive Office of the President
Washington, DC 20503
Ph: 202-395-4814

Workshop on Security Procedures for the Interchange of Electronic Documents

National Institute of Standards and Technology
Gaithersburg, Maryland
November 12-13, 1992

Abridged Version

LINKING SECURITY AND THE LAW OF COMPUTER-BASED COMMERCE

by

Michael S. Baum, J.D., M.B.A.
Independent Monitoring
Cambridge, Massachusetts USA

PREFACE

It is frequently (and astutely) stated that the law has not kept pace with technology. The historical tensions of law reform intended to accommodate technological change are manifested in the words of Oliver Wendell Holmes, who said

> [a]s few could write, most people had to authenticate a document in some other way, for instance, by making their mark. This was, in fact, the universal practice in England until the introduction of Norman customs. With them seals came in. But as late as Henry II they were said by the Chief Justice of England to belong only to kings and to very great men. I know no ground for thinking that an authentic charter had any less effect at that time when not under seal then when it was sealed. . . . Its conclusive effect was due to the satisfactory nature of the evidence, not to the seal. . . . But when seals came into use they obviously made the evidence of the charter better, in so far as the seal was more difficult to forge than a stroke of the pen.[1]

Similarly, the Supreme Court stated that

> [f]ormerly wax was the most convenient, and the only material used to receive and retain the impression of a seal. . . . We cannot perceive why paper, if it have that capacity, would not as well be included in the category. The simple and powerful machine, now used to impress public seals, does not require any soft or adhesive substance to receive or retain their impression. The impression made by such a power on paper is as well defined, as durable, and less likely to be destroyed or defaced by vermin, accident, or intention than that made on wax. It is the seal which authenticates, and not the substance on which it is impressed; and where the court can recognize its identity, they should not be called upon to analyze the material which exhibits it.[2]

Just as prior generations have grappled with document trustworthiness, today we must creatively forge a path which accommodates current requirements and practices, while contemplating the future. *Solutions* necessarily require compromises -- the challenge is to develop solutions and compromises that are thoughtful, practical and extensible. This is a daunting undertaking, but it is, at the same time, necessary and exciting.

---

1 OLIVER WENDELL HOLMES, THE COMMON LAW 272-273 (1881).

2 Pillow v. Roberts, 54 U.S. (13 How.) 472, 473-74 (1851).

TABLE OF CONTENTS

ACKNOWLEDGMENTS

The author gratefully acknowledges the comments and suggestions of many people, and particularly the significant comments and suggestions of the following people: Thomas Armstrong, Esq., U.S. General Accounting Office; George Chandler, Esq., Hill, Rivkin, Lomberg, O'Brien, Mulroy & Hayden; Douglas S. Cohen, Boston University Law School; Jerry Cohen, Esq., Perkins, Smith & Cohen; Clyde Christofferson, Esq.; Richard Dodd, Esq.; Sandy Epstein, Racal-Guardata, Inc.; Robert Fougner, Esq., PKP; Françoise Gilbert, Esq.; Altheimer & Gray; Gregory A. Gilbert, Boston University Law School; Ted Humphreys, XISEC Consultants Ltd.; Claire Johnson, Esq., Wilde Sapte; Gregory P. Joseph, Esq., Fried, Frank, *et al.*; Steve Kent, Ph.D., BBN; Professor Emeritus Alfred I. Maleson, Suffolk University Law School; Jerry Rainville, Esq., NSA; Miles Smid, NIST; Thomas Smedinghoff, Esq., McBride, Baker & Coles; Lee Stanton, General Electric Information Services; Oliver Smoot, Esq., CBEMA; and Peter Weiss, Esq., OMB.

AUTHORSHIP

Michael S. Baum is Principal of Independent Monitoring, a Cambridge, Massachusetts consultancy specializing in electronic data interchange and electronic commerce law and security. Baum chairs the EDI and Information Technology Division and the Information Security Committee, Section of Science and Technology, American Bar Association. The views expressed in this article do not necessarily reflect those of any organization or person other than the author. Because this paper presents some new or otherwise untested ideas, and because the subject matter of this paper begs further debate and consideration, comments and criticism are respectfully solicited.

Michael S. Baum
33 Tremont Street
Cambridge, MA 02139-1227 USA
FON: 1-617-661-1234
FAX: 1-617-661-0716
INTERNET: baum@hulaw1.harvard.edu

LINKING SECURITY AND THE LAW OF COMPUTER-BASED COMMERCE

by

Michael S. Baum, J.D., M.B.A.

I. INTRODUCTION

The accelerating movement from paper-based transactions and records to their electronic replacements, and the resulting benefits from this movement, are well documented. Yet in many cases, the shift from conventional to electronic mechanisms has not enjoyed sufficient legal consideration and treatment. Real and *perceived*[1] security weaknesses of electronic transactions and records remain legal and practical barriers to their effective widespread use. This paper considers the legal efficacy[2] and expanded use of electronic transactions and records in modern commerce, government, and other environments for undertaking commitments and other important purposes. The paper also asserts that information security mechanisms exist, considers their associated costs and benefits, and advocates, where appropriate, the use of such mechanisms. A model security baseline is proposed. The goal is to arrive at a reasonable level of security for various classes of transactions and records to provide assurances of satisfying legal requirements. The thrust of this paper, however, focuses on the legal implications of authentication, integrity, non-repudiation and availability rather than on those of confidentiality. This focus is not intended, however, to underplay the criticality of responsive private and government treatment of confidentiality issues -- indeed, confidentiality is the most critical requirement in some applications.[3] While this paper presents some "action-oriented" proposals, clearly the work has only begun.

---

1 Arguably, perceived security weaknesses could be reduced or eliminated by accepting commercially reasonable security practices (*see infra*). The failure to do so causes perceived weaknesses to become unnecessary barriers.

2 *Legal efficacy* in this paper denotes wide legislative and judicial recognition that properly secured electronic transactions and records satisfy traditional legal indicia of reliability. These indicia include, where appropriate, transactions or communications considered to be *in writing, signed, verified,* or *acknowledged.* Such legal requirements often differ considerably among states and among nations, as well as by application.

This paper neither endorses nor condemns *writing, signing,* or other requirements that historically support conventional paper-based attestations and commitments. Legal analysis of these requirements and responsive private and legislative reform should consider and reflect pragmatically the underlying attributes and objectives of such requirements (*e.g.,* authentication and integrity). A mere redefinition of *writing* and *signature* is not recommended.

3 *See, e.g.,* WORKGROUP FOR ELECTRONIC DATA INTERCHANGE (WEDI), REPORT TO SECRETARY OF U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES, Recommendation 8 (relating to confidentiality -- the WEDI recommendations did not include a comparable recommendation on information authenticity and integrity) (July 1992).

II. SECURITY AND RELIABILITY

a. Treatment in the Law

The creation, processing, communication, control, management, storage, use, retention, and retrieval of information in electronic form[4] have become critical to modern society. However, Electronic Data Interchange ("EDI") and transactions and records in electronic form are not yet accorded the extent of the legal efficacy enjoyed by paper-based transactions and records. Before these electronic forms can earn this legal efficacy, they must establish customs and practices, or they must at least be judged legally equivalent to their manual counterparts.[5] This problem of legal efficacy arises in the following areas of law, among others: contracts, evidence, government procurement and regulation, criminal law, real property, and the judicial process.

1. Contracts -- Seeking to satisfy requirements for electronic transactions and records under the Uniform Commercial Code ("U.C.C."), raises certain fundamental issues.[6] For example, although the definition of *signed* in U.C.C. § 1-201(39) "includes any symbol executed or adopted by a party with the present intention to *authenticate* a writing" (emphasis added), the word *authenticate* is not defined in U.C.C. Articles 1 or 2 (although Official Comment 39 to U.C.C. § 1-201 includes mention of a thumbprint (a particularly forensically-intensive[7] type of authentication). This lack of definition has created confusion in the legal community. While the case law considering electronic writings and signatures is sparse and inconsistent, some of those cases addressing the issue confirm the importance of the probative value of signatures.[8]

---

4 Hereinafter, references to *records* or *information in electronic form* will include their electronic creation, processing, communication, control, management, storage, use, retention and retrieval unless expressly qualified.

5 EDI technical and security standards do not serve as a substitute for responsive legal consideration. Such standards are purposefully drafted to provide options and alternatives to accommodate use by diverse industries and do not necessarily provide the guidance necessary to assure the creation of unequivocal legal acts. Technical standards developers cannot properly analyze and resolve complex legal issues.

6 *See, e.g.*, U.C.C. § 2-201 (Statute of Frauds); U.C.C. § 1-201(39) (defining "signed") and U.C.C. § 5-104 (addressing Formal Requirements and Signing).

7 *See generally*, BAUM, EDI AND THE LAW (Walden, ed. 1989) § 9.4 "The signing Requirement" at 123-125 (asserting that signatures and their electronic analogs should carry "forensic characteristics providing probative evidence of identity").

8 *See* In re Carlstrom, 3 U.C.C. Rep.Serv. 766, 773 (Bankr. D.Me. 1966) (requiring the affixed symbol for signature purposes under U.C.C. § 9-402 (Formal Requisites of Financing) to be *susceptible of evidentiary connection to the signatory*).

2. Evidence -- The Federal Rules of Evidence do not address specifically electronic digital data security mechanisms.[9] The scope of proof of trustworthiness (and, arguably, security) as an evidentiary foundation requires closer scrutiny. "[B]ecause electronic files are particularly susceptible to purposeful or accidental alteration, or incorrect processing, laying a foundation for their admission must be done with particular care. Proper control over creation and maintenance of these files can be crucial in overcoming inevitable objections that will be raised in the courtroom."[10] The implications of burgeoning, open, interconnected, and highly diverse computer systems utilizing expert system components, which may change frequently and considerably, may call for strong evidentiary foundations.[11]

There is some case law supporting the notion that proof of reliability (and implicitly security) is recognized as appropriate and necessary in evaluating the admissibility of computer-based evidence.[12] Other cases suggest a relaxation of the foundation required for admissibility of certain computer-based information (absent abuse of discretion by the judge).[13]

The Manual for Complex Litigation Second (1985) recognizes and addresses this problem of proof of reliability, yet by focusing on weight rather than admissibility, it reaches an equivocal, and ultimately unsatisfactory, solution of such evidentiary issues. It observes that "[n]otwithstanding the capacity of computers to make tabulations and calculations involving enormous quantities of information . . . several sources of potential errors of great magnitude exist."[14] The Manual further notes that the proponent of computerized evidence has the burden of laying a proper foundation by establishing its accuracy,"[15] and "the existence or possibility of errors usually affects only the weight, not the admissibility of the evidence, except when the problems are so significant as to call for exclusion . . ."[16]

---

9 *Cf.* FED. R. EVID. 901(b)(9) (Process or system), 1001(1) (Writings and recordings), 1001(3) (Original), 902 (Self-Authentication), and N.J. R. EVID. 1(13) (writing); *see* Peritz, *Computer Data and Reliability*, 80 Nw. U.L. Rev. 956 (1986) *reprinted in* 7 Comp. L.J. 23 (1986).

10 U.S. DEPT. OF JUSTICE, ADMISSIBILITY OF ELECTRONICALLY FILED FEDERAL RECORDS AS EVIDENCE: A GUIDELINE FOR FEDERAL MANAGERS AND COUNSEL (Oct. 1990) at 2.

11 *See* Section V. INTEGRATING FORMALISTIC AND EVIDENTIARY REQUIREMENTS, *infra* (examining evidentiary requirements for the laying of a foundation).

12 *See* U.S. v. Scholle, 553 F.2d 1109, 1124-25 (8th Cir.), *cert. denied*, 434 U.S. 940 (1977) (stating that computer storage needs a more comprehensive foundation for admissibility, including testimony on procedures for input control, such as a test for insuring accuracy and reliability).

13 *See, e.g.*, Rosenburg v. Collins, 624 F.2d 659 (5th Cir. 1980); U.S. v. Vela, 673 F.2d 86, *reh'g den.* 677 F.2d 113 (5th Cir. 1982), and U.S. v. Linn, 880 F.2d 209 (9th Cir. 1989). Note, however, that each of these cases involved telephone company billing records -- records which are created and retained by *trusted third parties*.

14 MANUAL FOR COMPLEX LITIGATION SECOND § 21.446 (1985).

15 *Id.*

16 "In view of the complex nature of the operation of computers and general lay unfamiliarity with their operation, courts have been cautioned to take special care to be certain that the foundation is sufficient to warrant a finding of trustworthiness and that the opposing party has full opportunity to

3. Government Procurement and Regulation -- Interpretation and resolution of State, Federal and foreign requirements such as those concerning signature requirements remains unsettled. Compare the following varied -- arguably conflicting -- signature definitions.

i. *signature* - "includes a mark when the person making the same intended it as such"[17];
ii. *signed* - "includes any symbol executed or adopted by a party with the present intention to authenticate a writing"[18];
iii. *signed* - "shall include the entry in the form of a magnetic impulse or other form of computer data compilation of any symbol or series of symbols executed, adopted or authorized as a signature"[19];
iv. *signature* - "in the case of an EDI transmission, means a discrete authenticating code intended to bind parties to the terms and conditions of a contract"[20]; and
v. *electronic signatures* - "characters representing the nominated persons on documents, and signed or symbols identifying their writers."[21]

One working group which considered this issue apprehended the effect of such uncertainty when it concluded that "[t]he lack of adoption of an accepted electronic signature policy by the [Department of Defense] will prevent some contract transactions being conducted in digital form."[22] Independently, the Comptroller General has addressed uncertainty in electronic commerce with the following decision: "[c]ontracts formed using Electronic Data Interchange technologies may constitute valid obligations of the government for purposes of 31 U.S.C. § 1501, so long as the technology used provides the same degree of assurance and certainty as traditional 'paper and ink' methods of contract formation."[23] Nevertheless, outside of the specific circumstances presented in the NIST case, the decision begs for a closer definition of the indicia of assurance and certainty necessary to be deemed reliable.

---

inquire into the process by which information is fed into the computer." MCCORMICK, HANDBOOK OF THE LAW OF EVIDENCE, (2d Ed. 1972) at 734.

17 1 U.S.C. § 1.

18 U.C.C. § 1-201(39).

19 17 C.F.R. § 230 (1990).

20 41 C.F.R. § 101-41.002(d) (1990).

21 Korean Act on Promotion of Trade Business Automation (1992) (Law No. 4479 Enacted Dec. 31, 1991) Art. 2.8 (Definitions, "Electronic Signature") reprinted in UN/ECE/TRADE/WP.4/R.872 (Aug. 4, 1992) (hereinafter "Korean Act") at 5.

22 Legal Issues Committee of the Acquisition Task Group, CALS/EC Industry Steering Group, Report on Potential Legal Issues Arising from the Implementation of CALS (Nov. 10, 1991) at 10.

23 Matter of National Institute of Standards and Technology--Use of Electronic Data Interchange Technology to Create Valid Obligations, Dec. of the Comp. Gen. of the U.S., File B-245714 (Dec. 13, 1991). *See* TABLE 2 - FALLIBILITIES OF PAPER-BASED SIGNATURES, *infra* Section.II.c.

4. Real Property – An example of how the problem of legal efficacy of electronic information could arise in the real property area involves the recording of deeds and related instruments where the recording statute mandates that "writings which are to be recorded or docketed in the clerk's office of courts of record in this Commonwealth shall be an original or first generation printed form, or legible copy thereof, pen and ink or typed ribbon copy. . . ."[24] Such a statute raises considerable barriers to computer-based commerce.

5. In Relation to the Judicial Processes -- The legal efficacy of information in electronic form also arises in judicial contexts. Despite the advance of computer automation in some aspects of the judicial process, electronic notice and service of process are not generally permitted by court rules. However, there are exceptions,[25] and judicial reform is accelerating.[26]

It is evident from the above discussion of the different legal fields that there is need for legal reform. As noted in the *Uniform Rules of Conduct for Interchange of Trade Data by Teletransmission* ("UNCID"), a pioneering international code of conduct that addresses important legal and control considerations attendant to the use of electronic trade data, the

> electronic document is quite different [from a paper document]. It takes the form of a magnetic medium whose data content can be changed at any time. Changes or additions will not appear as such . . . . it is possible to establish techniques which give electronic data interchange characteristics that make it equal or superior to paper not only as [a] carrier of information, but also as regards the evidential functions.[27]

Moreover, electronic transactions are increasingly communicated within open, distributed and interconnected environments.[28] These environments potentially expose users and networks to risks from both accidental and deliberate alteration and destruction of data,[29] because open environments are generally more

---

24 VA. CODE § 55-108.

25 *E.g.*, FED. R. APP. P. 25(a) (1991) (authorizing a court of appeals to accept papers filed "by facsimile or other electronic means"); OHIO R. C. P. Rules 5(e) and 10 (July 1, 1991). The National Archive and Records Administration's *Electronic Records Management* regulations accommodate the judicial use of electronic records pursuant to FED. R. EVID. 803(8). 36 C.F.R. § 1234.24 (1990).

26 Additionally, the U.S. Department of Justice has issued findings which "encourage the development of electronic data interchange technologies." BUREAU OF JUSTICE STATISTICS, REPORT OF THE NATIONAL TASK FORCE ON CRIMINAL HISTORY RECORD DISPOSITION REPORTING, NCJ-135836 (June 1992) at 1.

27 INTERNATIONAL CHAMBER OF COMMERCE, Pub. No. 452 (1988) at 8.

28 *See generally*, NATIONAL RESEARCH COUNCIL, COMPUTERS AT RISK, SAFE COMPUTING IN THE INFORMATION AGE (1991) (hereafter "NRC").

29 *See* Eckerson, *Network security lacking at major stock exchanges*, Network World (Sept. 16, 1991) at 23-24; Prefatory Note, U.C.C. Art. 4A (1990); *see also* Shell Pipeline v. Coastal States Trading, 788 S.W. 2d 837 (Tex. Ct. App. 1990) (Shell's responsibility for correction of errors was upheld, even where Shell's undertaking was "entirely gratuitous").

difficult to control than are closed ones.[30] "New vulnerabilities are emerging as computers become more common as components of domestic and international financial systems. *The nation needs computer technology that supports substantially increased safety, reliability, and, in particular, security.*"[31]

Additionally, in open environments, parties will increasingly desire or need to communicate and make commitments without having executed electronic trade and communication agreements. Consequently, the degree of *end-to-end* security[32] in such trading environments takes on increased importance.[33] "[I]f the information is shared between user groups or exchanged via a public or generally accessible [] network . . . [n]either the technology, terminals and services nor the related standards and procedures are generally available to provide comparable security for information systems in these cases."[34]

Although the extent (or strength) of the security necessary to support reliable electronic transactions and records for legal purposes is unclear, security is increasingly recognized as critical.[35] This conclusion is supported by decisions, studies and opinions of public and private entities. For example, the United Nations Commission on International Trade Law (UNCITRAL) stated that "it is clear that the legal reliability of EDI techniques requires that high standards be used to determine legal certainty as to the identity of the sender, its level of authorization and the integrity of the message."[36] The Comptroller General of the United States has remarked that "[a]gencies can create valid obligations using *properly secured* EDI systems."[37]

---

30 Because conventional management techniques and controls cannot respond adequately to open and distributed environments, technology-based techniques and controls may be necessary.

31 NRC, *supra* at 2 (emphasis by Council). This view is substantiated by reports of increasing problems. For example, "[i]t [was] estimated that security breaches, including lost revenue, data recovery, lost computing time, and personal downtime . . . cost U.S. corporations $1 billion in 1990." YANKEE GROUP, DATA NETWORK RELIABILITY AND SECURITY (1990).

32 End-to-end security refers to those sets of services that are applied to information prior to their submission to the communication mechanism. These services provide security assurances throughout the transfer to the intended recipient and which are verifiable by the recipient. Such services may include, but are not limited to, digital signatures for authenticity and integrity, and encryption for privacy purposes.

33 The U.S. Department of Defense has recognized the weaknesses in such open communications environments: "[i]t is important to reiterate that the CN [communication network] is not relied upon for the confidentiality or integrity of the information it transfers. Failures in a CN can only result in the delay, mis-delivery, or non-delivery of otherwise adequately protected information." Draft DOD INFORMATION SYSTEMS SECURITY POLICY at § 4.4 "FIRST PROTECTION ALLOCATIONS" (March 30, 1992) (note: this is not yet DoD policy).

34 E.C. *supra* at Annex, Action Line 3.1.

35 E.C., *supra* Action Line 4.1. ("In the security of information systems there is inherently a very close relationship between regulatory, operational, administrative and technical aspects.").

36 *Electronic Data Interchange*, Rep. of the Sec. Gen., UNCITRAL, 246th Session, Vienna, 10-28 June, 1991, U.N. Doc. A/CN.9/350 (15 May 1991) at 23.

37 Dec. of the Comp. Gen., *supra* (emphasis added).

Other supporting opinions can be seen in model trade agreements and the developing literature. A model EDI agreement states that "[a]dequate security procedures are recognized. . . as critical to the efficacy of electronic communication. . . . The use of adequate security enhances the reliability of those records and enhances the ability to prove the substantive terms of any underlying commercial transaction."[38] A further supporting view notes that *"[l]egal reliability* actually implies 'demonstrably and unarguably high standards of authorization, [sic] operational and access control and management' use of IACT [information and communication technology] systems. 'Authorisation,' further, implies 'accurate, precise and dependable identification, verification and authentication technologies and techniques which are, or may become, as legally acceptable as the conventional trust and comfort of a manual signature written in ink on paper."[39]

b. Reasonable Security Procedures

Unlike conventional paper-based transactions and records, there is little jurisprudential guidance as to whether (and, if so, under what circumstances) a particular security technique, procedure or practice will provide the requisite assurance of reliability in electronic form. This lack of guidance concerning security is reflected in the multiplicity of current security and authentication practices within the EDI community. These practices, in many instances, appear to have been implemented in an *ad hoc* manner, with neither a clear understanding of the present state of law, nor the technical proof assurances of other chosen practices.[40] Where the law has responded, it has been arguably too vague -- such as a requirement to implement *reasonable security procedures.*[41]

> While security procedures should certainly be reasonable, in certain situations a lack of specificity in defining "reasonable" security procedures may provide inadequate guidance causing such security procedures to fail in their intended purpose. . . . Specificity may help the parties implement and comply decisively and unambiguously with security procedures, reduce confusion and offer better expectations of reliability and certainty.

---

38 A MODEL EDI TRADING PARTNER AGREEMENT § 1.4 Comment 1, *supra.*

39 Stephen Castell, *"The Legal Admissibility of Computer Generated Evidence Towards 'Legally Reliable' Information and Communications Technology (IACT),"* COMP. LAW AND SEC. REP. (Jul.-Aug. 1989) at 7-8. (discussing the Appeal Study *Appendix on Evidence Admissible in Law* by S. Castell and the Central Computer and Telecommunication Agency, British Treasury, 1988; subsequently published as The Appeal Report, May, 1990).

40 In a survey of EDI users, the mechanisms or procedures employed as legal signatures included the following: a "buyer code," a DUNS number and suffix, a password, a message authentication code, an account number, an ID/password combination, an "electronic verification of symbol and codes," and functional acknowledgments. LEGAL AND BUSINESS CONTROLS TASK GROUP, ACCREDITED STANDARDS COMMITTEE X12, 1990 SURVEY (1990).

41 For example, in banking, a *security procedure* has been defined as: "a procedure established by agreement of a customer and a receiving bank for the purpose of (i) verifying that a payment order or communication amending or canceling a payment order is that of the customer, or (ii) detecting error in the transmission or the content of the payment order or communication." U.C.C. § 4A-201 "Security Procedure."

> Security procedures should be sufficiently precise so that they are not subject to discretionary, self-serving interpretation, in part, because: (i) few standard security procedures exist in the law. . . . (ii) security technology is changing rapidly, and (iii) parties often hold particularly diverse opinions on appropriate solutions to security threats.[42]

One difficulty in developing responsive laws involves deciding the extent to which law should detail and endorse particular security techniques, procedures or practices.[43] Proponents of specificity argue that the electronic commerce community needs greater guidance[44] and that private agreements and legislation requiring only *reasonable security procedures* are vague and unworkable. Proponents of generality, on the other hand, argue that the endorsement of specific security procedures, practices or techniques leads to inflexibility and creates a presumption that the failure to implement such techniques, procedures and practices constitutes failure to exercise ordinary care. While recognizing these competing interests, a stronger viewpoint supports a measured movement toward greater specificity.

The electronic commerce community is asking lawyers to consider and to provide advice concerning signatures, security procedures and other related issues, but as yet, the legal community's experience with these issues is limited. Attorneys often defer to security professionals, who in turn seek the guidance of auditors, who then defer to attorneys. This *circle of deference* suggests that sufficiently concise answers to, responsibility for, and the resolution of, these issues are not quickly forthcoming. Moreover, it suggests that there is need for professional education in the system.[45] Further study is warranted in this area. Lawyers, security professionals and auditors should strive to provide education as a means to develop ideas on what attributes reasonable security would possess, as well as to identify responsive security services, their associated strengths, and when they can and should be implemented.

Consistent with this approach, the House of Delegates of the American Bar Association (ABA) has approved the first ABA Resolution that directly responds to critical legal-security issues affecting electronic data interchange and electronic commerce. The resolution requires the ABA to do the following:

---

42 MODEL ELECTRONIC PAYMENTS AGREEMENT AND COMMENTARY, § 7 Comment 5, (June, 1992), prepared by the EDI and Information Technology Division, Section of Science and Technology, American Bar Association, (hereinafter "MODEL AGREEMENT") 32 JURIMETRICS J. No. 4 at 601 *et seq.* (1992)); *see generally*, Michael S. Baum, *Commercially Reasonable Security: A Key to EDI Enforceability*, ACTIONLINE, (AIAG, Nov. 1989).

43 Various legislation and guidelines mentions, or recognizes specific security technologies.

44 "Buyers of cryptography cannot independently evaluate a seller's claims of product security." Sandy Epstein, "Striking a Balance: View on a National Cryptographic Policy," testimony before the National Computer System Security and Privacy Advisory Board, NIST (Gaithersburg, Sept. 1992).

45 There are only a few formal law school course offerings applicable to computers and EDI legal issues and course offerings on information security legal issues are probably nonexistent. "A lack of EDI education is perhaps today's greatest hindrance to productive EDI usage and such implementation." Sokol, *EDI Education Pays Dividends*, Data Interchange (Dec. 1991) at 16.

[s]upport action by federal and state governments, international organizations, and private entities to:

a) facilitate and promote the orderly development of legal standards to encourage use of information in electronic form, including appropriate legal and professional education;

b) encourage the use of appropriate and properly implemented security techniques, procedures and practices to assure the authenticity and integrity of information in electronic form; and

c) recognize that information in electronic form, where appropriate, may be considered to satisfy legal requirements regarding a writing or signature to the same extent as information on paper or in other conventional forms *when appropriate security techniques, practices, and procedures have been adopted.*[46]

Consistent with the ABA approach, the United Nations Commission on International Trade Law ("UNCITRAL"), as early as 1985, recommended that governments "review legal requirements of a handwritten signature or other paper-based method of authentication on trade related documents with a view to permitting, where appropriate, the use of electronic means of authentication."[47]

c. Mapping Security Attributes to Legal Standards

There are various techniques available, with specified assurances to authenticate the source of, verify the content of, and control access to, data in electronic form. Many more of these techniques will develop as both the technology and the law evolve. History has demonstrated repeatedly that legal rules prescribing technology for authentication and related purposes have been a function of the available technology, historical accident or anomaly, and the technology's forensic[48] characteristics. It has also been transitory.[49]

The following table (TABLE 1) presents some of the attributes of conventional writing and signings as compared to their *approximate* electronic security analogs. The strength (and the propriety of the suggested analog) of any such security mechanism depends considerably on its implementation and the associated system controls. For example, in the case of a "signature" requirement, any appropriate

---

46 Developed and submitted by the Section of Science and Technology to the House of Delegates of the ABA, the Resolution (no. 115) was approved on Aug. 19, 1992 (emphasis added).

47 OFFICIAL RECORDS OF THE GENERAL ASSEMBLY, FORTIETH SESSION, SUPPLEMENT NO. 17 (A/40/17), ¶ 360.

48 *See generally*, BAUM, EDI AND THE LAW, *supra* § 9.4 "The Signing Requirement" at 123-125 (asserting that signatures and their electronic analogs should carry "forensic characteristics providing probative evidence of identity").

49 *See* Preface, *supra* (providing quotes that give insight into the forensic and transitory nature of technology-related rules).

security technique that provides comparable or superior attributes to those produced by the conventional use of a written signature should be satisfactory.[50] However, the various security attributes in TABLE 1 demonstrate that the handwritten signature does not have an unequivocal electronic analog.[51]

---

50 Conventional paper-based handwritten signatures inherently have security attributes to the extent that, *e.g.*, ink cannot easily be erased without detection, paper is non-transient, and a signature is biometrically unique.

51 These three examples of information in electronic form (categories "B," "C" and "D" in TABLE 1) are also used to support the security services provided in the Model Security Baseline in Section III.c., *infra*.

Attribute	A Conventional Signed Writing Communicated via United States Postal Service	B Unencrypted Information with Symbol in Electronic Form Communicated via Third party Service Provider	C Digitally Signed or Cosigned Information in Electronic Form [52] Communicated via Third party Service Provider	D Digitally Signed & Notarized Information in Electronic Form Communicated via Third party Service Provider
Origin Authen.	Medium-Strong	Weak	Strong	Strong
Proof of Receipt	Return Receipt	Weak	Strong	Strong
Content Integrity	Partial	Weak	Strong	Strong
Time of Creation	Weak	Weak	Weak	Strong
Time of Dispatch	Postmark	Weak	Weak	Strong
Time of Receipt	Return Receipt	Weak	Weak	Strong
Time of Acknow.	Return Receipt	Weak	Weak	Strong
Singularity	Yes	No	No	Can be offered as a "registry" service
Biometric	Yes, signature	No, but available for resource access control	No, but available for resource access control and for cryptoignition[53]	No, but available for resource access control and for cryptoignition
Expression of Intent[54]	Indicia	Indicia	Indicia	Indicia
Non-repudiation	Partial	Weak	Strong, except time	Strong
Privacy	If enveloped[55]	Weak	Weak	Weak

TABLE 1 - COMPARISON OF SIGNED WRITINGS AND ELECTRONIC INFORMATION*

Key to TABLE 1

* General Comment: Attributes exhibiting a propensity for forgery are listed as "Weak." TABLE 1 includes subjective positions and is intended exclusively for pedagogical purposes.

A: A signed paper document sent by postal service.

B: Unencrypted (clear text) communicated via third party service provider (TPSP). Satisfaction of many listed attributes depends largely on controls, including TPSP controls.

C: Digitally signed electronic document.

D: Digitally signed electronic document which is "notarized" (time stamped and digitally signed) by a trusted third party. In this Table, notarization is available (via *trusted box*) at the site of origin, at the respective TPSPs and at the site of receipt.

---

52 While many security services are best implemented using digital signatures or comparable cryptographic methods, many can be implemented non-cryptographically, although not necessarily with comparable economy, strength, functionality or elegance.

53 A quantity which enables a cryptographic algorithm(s) or device embodying a cryptographic algorithm(s) to operate which is generally implemented as a component of a secret quantity used to convert other quantities necessary for operation.

54 This may vary among criminal and civil proceedings.

55 "The Postal Service must preserve and protect the security of all mail in its custody from unauthorized opening, inspection, or reading of contents or covers, tampering, delay or other unauthorized acts." DOMESTIC MAIL MANUAL (DMM) § 115.1 "Importance of Mail Security;" "In general, no person may open, read, search, or divulge the contents of mail sealed against inspection . . ."

One additional comparison is instructive. A decision of the Comptroller General proffered three signature attributes as being necessary to create obligations which can be recorded against the government. TABLE 2 considers these attributes within the context of fallibilities of paper-based media.[56]

PROPOSED SIGNATURE ATTRIBUTES[57]	FALLIBILITIES
Unique to the Certifying Officer	Forgery. Where stamps and other mechanisms are used, the signature is not unique to the certifying officer.
Capable of Verification	Error prone. Signature comparison is an art as well as a science; verification often disregarded due to cost, ineffectiveness or unavailability.
Under Officer's Sole Control	Law permits other mechanisms which may not, without knowledge of custom and practice, provide assurances of sole control.
Proposed effect:	
Demonstration of Intent to be Bound[58]	Depends on the circumstances of its use. Not an inherent attribute.

TABLE 2 - FALLIBILITIES OF PAPER-BASED SIGNATURES

To the extent, *arguendo,* that the Comptroller General's decision is interpreted to substantially require the use of cryptographic methods[59], three observations deserve consideration. First, despite an inference that paper-based signatures provide a good benchmark for authentication and provability, Table 2's proffered signature fallibilities effectively present a compelling case that supports the permissibility of non-crytographically enhanced transactions where appropriate.[60] Second, the noted weaknesses of conventional signatures relative to digital signatures (*see* TABLE 1) support the legal efficacy of digital signatures in substitution for the latter. Third, although the decision does not expressly reference

---

56 *Cf.*, the quotes in the Preface to this paper concerning fallibilities of conventional media.

57 Proposed by the Comptroller General of the United States.

Other signature attributes which have been proposed within the private and commercial sectors include attributes that: *identify* the signatory to the transaction; *demonstrate* that the signatory had the intent to formalize the information due to its importance; *create* a record acceptable to the dispute resolution mechanism; *evidence* the existence of a contract; and, *prevent* repudiation.

58 This is not a formal attribute but instead a conclusion. Note also that some government representatives advocate that having established a signature, it is also necessary to demonstrate that the signature is *linked to the data.*

59 Cryptography "embodies principles, means and methods for the transformation of data in order to hide its information content, prevent its undetected modification and/or prevent its unauthorized use." ISO 7498-2-1988(E) § 3.3.20.

In further support of the appropriate use of technology-based solutions, *see* FINANCIAL MANAGEMENT SYSTEMS, OMB Circular No. A-127 (obliging government agencies to *use the most contemporary technology*); and T. J. Hooper, 60 F.2d 737 (1932) ("[a] whole calling may have unduly lagged in the adoption of new and available devices").

60 *See infra* Section III.c. "A Model Security Baseline."

non-repudiation, it effectively focuses on the attributes of non-repudiation, thereby bolstering the utility of this comparatively unfamiliar service.

d. Non-repudiation

Some security services can provide diverse capabilities that are not necessarily provided by conventional paper-based techniques. One such security service is known as a *non-repudiation service*. Generally, non-repudiation services prevent a document's originator from denying the document's origin and provide *irrevocable proof of authenticity*.[61]

> The Non-repudiation Service may be provided through the use of mechanisms such as digital signatures, encipherment, data integrity and notarization, with support from other system services such as Time Stamping and Security Services. The Non-repudiation Service can use a combination of these mechanisms and services as appropriate to satisfy the security requirements of the application in question. The goal of the service is to collect, maintain, make available and validate non-deniable proofs regarding data transfers between the originator and recipient.[62]

A non-repudiation service is presented as *one* example of a security service, which, whether or not cryptographically based, may satisfy requirements that are linked to conventional writings and signings, such as contributing to evidence of a party's intent to contract or to be bound. Although many existing legal requirements do not require absolute or non-repudiable proof, these security services offer the legal and control communities important tools and possibilities with which to fashion legal obligations to accommodate electronic practices (particularly the more important or risky obligations).

The time of the creation of a transaction or the submission of a transaction to an electronic messaging system, or the time when received or retransmitted by a third party service provider (TPSP), available to, received by, or acted upon by the intended recipient is critical in various applications. For example, where parties must file information electronically[63] (*e.g.*, tax returns), or where an electronic bidding process closes at a time certain, or where the first to file a response wins[64];

---

61 MESSAGE HANDLING: EDI MESSAGING SERVICE, CCITT Draft Rec. F.435 (Version 5.0, June 15, 1990).

62 ISO/IEC JTC1/SC21, Intro., WORKING DRAFT NON-REPUDIATION FRAMEWORK, N7082, Project 97.21.49.6 Q53 (July 1992).

63 The definition of *filing* has come under review. [insert references and relation to receipt and model agreements.]. "The word *file* is derived from the Latin work "filum,' and relates to an ancient practice of placing papers on a thread or wire for safe-keeping and ready reference. *See* MICHAEL S. BAUM AND HENRY H. PERRITT, JR., ELECTRONIC CONTRACTING, PUBLISHING AND EDI LAW (Wiley, 1991) [hereinafter "Baum and Perritt"] at § 5.16 "UCC Security Interest Filings" (considering many electronic filing issues).

64 *See* Abourezk v. Federal Power Commission, 513 F.2d 504, 505 (D.C. Cir. 1975) (where Judge Bazelon noted that "[d]ue to lack of synchronization between the clocks in the clerks' offices and those in the

trusted time stamping is recognized as a prerequisite to the proof of the completion of obligations of one party, and the transfer of obligations to another.

Despite the great benefits enuring from the use of digital signatures, they have some inherent limitations (as is true with any security mechanism) including an innate inability to provide "time-related" non-repudiation. Digital signatures and other cryptographic methods cannot, in the absence of a trusted entity, provide an unforgeable trusted time stamp. Therefore, to achieve *full* non-repudiation, time stamping must be undertaken by a disinterested party beyond the control of the parties to a transaction or record. Such a third party is a trusted entity.

e. Trusted Entities and Time Stamping

A trusted entity is an independent, unbiased entity capable of providing important security assurances that enhance the enforceability and reliability of electronic records. The key attributes of a trusted entity are that it is a *disinterested, unbiased, third party* trusted by the parties to the transaction and by the dispute resolution mechanism(s) relevant to a transaction or record. Simply stated, a trusted entity's administrative, legal, operational and technical infrastructure must be beyond reproach.[65]

A trusted entity can time and date stamp,[66] store (or forward) a "record copy" or hash of a transaction, keep an audited data log, or serve as an intermediary for other trust-based services between trading partners.[67] The trusted entity's record copy of an electronic transaction would control in the event of a dispute regarding a document's authenticity or timeliness.

The electronic notary[68] offers unique solutions to one of the critical "missing links" of electronic transactions and records assurances: unforgeable trusted time stamping. The electronic notary also may facilitate future TPSP and value added network service requirements by providing them with trusted-entity services.[69] The

---

offices of the various federal agencies, it is often not possible to be certain which petition was the first to be filed after the agency entered its order.").

65 Third Party Service Providers or value added networks, such as ATT or MCI (collectively "VANs") have arguably been inaccurately identified as trusted entities. VANs are not necessarily disinterested because they may compete with each other, participate in the transfer or processing of information (and therefore have exposure), and may introduce error, delay, unavailability or misdelivery.

66 The author offers a French term, which more concisely describes the intended time stamp functionality: *horodatage* (horo=hour, and datage=date). The use and significance of time stamping has both a rich historical as well as contemporary value.

67 *See* BAUM AND PERRITT, *supra* at Ch. 5 (providing an extensive survey of possible trusted entity - clearinghouse services).

68 The terms "notary" or "notarization" in the context of electronic transactions do not have recognized legal standing equivalent to that of the conventional notary public, and consequently, such terminology used in this setting is inaccurate or potentially confusing. *See* BAUM AND PERRITT, *supra* §§ 4.33-4.36 (presenting a survey of issues pertinent to the automation of the notary public).

69 Because the electronic notary is not controlled by TPSPs or VANs, reliance by users need not be placed exclusively on the internal controls of the TPSPs and VANs, except for availability.

electronic notary can provide irrefutable proof of the time of the origination of the document.[70] Notarizing data intended for record retention and archiving provides an unforgeable seal which may contain a time stamp and digital signature, together with additional audit, legal and security information intended to enhance its legal efficacy.[71]

III. RISK ANALYSIS AND RISK-BASED APPROACHES

a. Risk Analysis

To the extent that various methods to assure that reasonable security procedures have been considered and implemented in both the private and public sectors, results have been inconsistent -- just as attempts to satisfy amorphous requirements for *commercially reasonable security* have produced varying results.[72] Such inconsistent results are explained, in part, by the insufficient and varying analytical tools used to evaluate security requirements (and their legal efficacy), such as *risk analysis*.

> 'Risk analysis' is a procedure used to estimate potential losses that may result from system vulnerabilities and the damage from the occurrence of certain threats. Risk analysis identifies not only critical assets [and processes[73]] that must be protected but considers the environment in which these assets are stored and processed. The ultimate purpose of risk analysis is to help in the selection of cost-effective safeguards that will reduce risks to an acceptable level.[74]

The National Institute of Standards and Technology ("NIST") noted the need for EDI risk analysis in March 1991 when it required agencies to *employ risk management techniques*. Yet, NIST did not provide specific guidance on EDI risk

---

70 *Cf.*, the proofs available from an electronic notary to those available from the U.S.P.S.. For example, consider that "[m]ail deposited in a collection box or post office may, with proper identification, be recalled by the sender." DMM *supra* at § 152.71 "Who May Recall Mail."

71 *See supra* (WORKING DRAFT NON-REPUDIATION FRAMEWORK).

72 *See supra* Section II.b., Reasonable Security Procedures.

73 *See* Thomas A. Stewart, "The Search for the Organization of Tomorrow," Fortune, (May 18, 1992) at 94-94 (includes a proposal for viewing the organization horizontally by core processes -- each core process is a set of functions necessary to meet a major external objective such as inventory turnover or on-time delivery.

74 IRENE GILBERT, GUIDE FOR SELECTING AUTOMATED RISK ANALYSIS TOOLS, NIST Special Pub. 500-174 (1989) at 3. THE COMPUTER SECURITY ACT OF 1987, 40 U.S.C. § 759 note, P.L. 100-235 (1987), requires applicable federal agencies to develop a computer security and privacy plan "that is commensurate with the risk and magnitude of the harm resulting from the loss, misuse, or unauthorized access to or modification of the information contained in" each Federal computer system.

analysis.[75] The creation and enforcement of legal commitments undertaken electronically may require new criteria (such as EDI-relevant legal criteria) and approaches to risk analysis that have either not been developed or widely adopted.[76] For example, EDI may involve variables and *higher order effects* that are difficult to quantify and that effectively require consideration of the legal interrelationship between a series of related EDI transactions and records without direct conventional analogs.[77] "EDI/EFT is too young for its full risk implications to become apparent."[78] There should be a move toward the development of authoritative risk analysis for electronic commerce in both the private and public sectors.

b. Security Baseline Issues

In considering various approaches to linking technical security measures and the law, it is important to recognize that the strength and reasonableness of security procedures for particular applications are risk driven. These procedures, therefore, must undergo further scrutiny. A *security baseline*[79] ("baseline") is a tool that may help define and rationalize security requirements for diverse electronic transactions and records. A baseline serves as a foundation to develop a clear expression of security requirements, facilitate open trading environments, ensure that transaction costs are commensurate with the risks, and provide greater legal certainty.[80]

A baseline can encompass generally accepted security methods and procedures (to the extent available to attain reasonable security at the operating system, data communication, and application (including EDI) levels).[81] Compliance with such requirements would establish a presumption of the security procedure's sufficiency

---

75 NIST, ELECTRONIC DATA INTERCHANGE (EDI), FIPS-PUB 161, 56 Fed. Reg. 13,123 (Mar. 29, 1991). *Cf.*, NIST COMPUTER SYSTEMS LABORATORY (CSL) BULLETIN, SECURITY ISSUES IN THE USE OF EDI (June, 1991).

76 Existing risk analysis tools focus neither on legal requirements nor on the particular needs of EDI. *See* NIST, GUIDELINE FOR AUTOMATED DATA PROCESSING RISK ANALYSIS, FIPS PUB 65 (Aug. 1979). *Cf.*, Birch and McEvoy, "A Structured Approach to Information Security Risk," COMP. LAW & SEC. REP., Vol. 8 Issue 4 (Jul.-Aug. 1992) at 177.

77 *E.g.*, EDI Functional Acknowledgment and Application Advice transaction sets do not exist in conventional paper-based practices. The loss or garbling of such transaction sets present challenges to conventional risk analysis. *See* BAUM AND PERRITT, *supra* at 180-181.

78 David Davies, "EDI Insurance - The 'Red Herring' Theory Examined," CLSR, Vol. 8, Issue 5 (Sept.-Oct. 1992) at 226-229 (noting "the relatively unproven or un-demonstrated nature of the risks;" and that "very little reliance should be placed upon the ability of existing insurances to encompass the new risks of EDI").

79 *See* Baum, Actionline, *supra* at 35 (advocating a security baseline).

80 The approach to the development of a baseline should be examined cautiously, considering that "[t]he law embodies the story of a nation's development through many centuries, and it cannot be dealt with as if it contained only the axioms and corollaries of a book of mathematics." OLIVER WENDELL HOLMES, THE COMMON LAW *supra*; and also taking into consideration that we should not take too *formulaic* an approach.

81 *See* NRC, *Recommendation 1 Promulgate Comprehensive Generally Accepted Security System Principles (GSSP) supra* at 27-32.

or legal efficacy.[82] A Baseline can take various forms, including legislation, private agreement and guidelines.[83]

The purpose of a baseline is to serve as a bridge between high-level policy and philosophical positions on one end of the spectrum, and, at the other extreme, detailed application-specific rules. As such, a baseline is positioned at an intermediate level of abstraction which both seeks to enforce high-level policy and provides a mechanism for the development of workable rules. TABLE 3 provides one perspective on how a baseline can be used and where it fits into the legal-standards environment.

ABSTRACTNESS	TYPE	COMMENTS
High	"Reasonable Security Procedures"	Too uncertain for rules; Preferably a policy objective
Medium	**Baseline** (single or multilevel)	A tool to help enforce policy objectives
Low	Application-specific rules and guidelines	Compliant with Baseline

TABLE 3 - RELATIVE LEVELS OF ABSTRACTION

Baseline security requirements should vary depending on risks and on other factors.[84] For low risk transactions -- *such as* those with a low probability of large losses, the benefits of strong security are likely outweighed by the costs of such measures.[85] Higher risk transactions may require more stringent controls, including cryptographic methods or trusted entity services.

A baseline should be sufficiently concise without regard to risk. The more specific the baseline, the greater will be the transactional certainty, user confidence, and ultimate success. Without specificity, security requirements may provide inadequate guidance and may fail in their intended purpose.[86] Specificity helps users to implement decisively and to comply unambiguously with baseline requirements. Consequently, until the parameters of *reasonable security* practices in electronic commerce become more clearly defined (as a function of improved experience and practice coupled with the use of better risk analysis tools), greater specificity is advocated. A baseline arguably fills this gap. Thereafter, generalized or abstract

---

82 *See infra* Section IV. BURDEN OF PROOF AND PRESUMPTIONS (presenting an alternative to the legal effect of the Model Baseline in TABLE 6).

83 *See* BAUM AND PERRITT, *supra* at 80-81 (discussing various forms of implementation guidelines).

84 See the various factors described later in this section.

85 For most electronically communicated commercial non-financial transactions, the security regime is typically little more than that provided by simple password/ID-based access or authentication controls. Such weak security probably results from established customs and practices, simplicity, lack of security sophistication, financial constraints, and the belief that password/ID-based access controls are the lowest common denominator (and, in this respect, are most pragmatic) for ubiquitous computer-based communications.

86 *See* Michael S. Baum, *Commercially Reasonable Security: A Key to EDI Enforceability*, Actionline, *supra; see also* BAUM AND PERRITT, *supra* at 184.

standards of "reasonable security" may become legally sufficient -- in an environment benefiting from legal precedent, and widely recognized specific procedures and practices.[87]

In order to understand this idea, the following issues should be considered in developing a baseline (but not necessarily be limited to):

1. Attribute-based Security Requirements - The security of particular transaction types, and of a particular transaction, will depend upon the needed security services and will vary as a function of risk and legal needs. Security services include authentication, integrity, non-repudiation, confidentiality, and availability.[88] TABLE 1, *supra*, presents security attributes within the context of a comparison to paper-based mechanisms.

2. Value Requirements - Developing consensus on a definition of value[89] is becoming a focal point in the development of electronic commerce rules[90] in both the private[91] and public sectors[92], and may go to the heart of the debate. The challenge is to determine which transactions, other than payment orders and "purely" financial transactions merit stronger information security protection (such as cryptographic-based authentication methods) than the security utilized for low value and low risk transactions. Two competing approaches on this issue are characterized as *narrow* and *broad*.

   •Narrow View Argument - Based on value, transactions which merit stronger information security are comparatively few in number. Generally, the narrow view does not provide increased protection for

---

87 Specific practices leading the way for acceptance of the general practice is analogous to traditional analysis in evidence law: initial close scrutiny of the trustworthiness of new technology prepares the way for future lenient acceptance of the new procedures -- coupled with a better understanding of the risks.

88 *See supra* (providing definitions for each of these security services).

89 *Value* may be defined broadly by the courts -- *e.g.*, as "[a]ny consideration sufficient to support a simple contract." Fowler v. Smith, 156 N.E. 913, 914 (Ohio App. ___). *Cf.*, U.C.C. § 1-201(44) (defining value broadly); U.C.C. § 2-714(2) ("Buyer's Damages for Breach in Regard to Accepted Goods);" and U.C.C. § 1-106 ("Remedies to Be Liberally Administered") (generally providing a subjective measure of damages; Official Comment 1 "rejects any doctrine that damages must be calculable with mathematical accuracy").

90 This includes the development of a Model Security Baseline, *see infra* Section IIc.

91 The National Automated Clearinghouse Association ("NACHA") does not distinguish between low and high value transactions where it "recommends that ACH [automated clearing house] processors and all ACH participants employ data security techniques in accordance with ANSI standards for authentication and key management." 1992 ACH Rules at OR xvii.

92 *Cf.*, "'Value' . . . will be determined on a case-by-case basis. In fact, Treasury itself moves very few funds. . . . The Treasury Directive on Electronic Funds and Securities Transfer Policy . . . makes it Treasury policy that *all* Federal EFT transactions be 'properly authenticated'. The authentication measures adopted . . . are those recommended by the American National Standards Institute (ANSI) in Standard X9.9." Treasury Directive 81-80, § 2.1.

purchase orders, "merely executory contracts,"[93] contract-related business documents (excluding payment orders) and other *low value*, low risk transactions. This view argues that business knows how to take care of itself, and business practices demonstrate that non-payment-related transactions are typically not communicated in a highly secured manner.[94] Business has determined that the cost of strongly securing purchase orders, invoices, and the like, is not commensurate with the risk.

•Broad View Argument - Individuals supporting this position believe that the scope of transactions of a value which merit stronger information security are comparatively broad. Many purchase orders, purchase order acceptances, and other non-payment documents require stronger security protection whenever the risk of loss or error associated with such documents threatens business assets or competitive position.[95] The value of a loss or error in a non-financial instrument may not necessarily result in as immediate a loss as with a financial instrument. However, the loss of, or litigation concerning, a non-financial instrument is nonetheless of comparable or greater value (such as where consequential damages are considered). Fiduciary duties owed by corporate management to its stock holders, include prudently protecting corporate assets -- and strong security is one prudent approach. Finally, paper-based practices demonstrate that the strength of security techniques implemented for low and medium value transactions do not vary considerably (except, *e.g.*, with respect to the use of multiple signatures for authorization), because low value transactions often are "bootstrapped" to a stronger security level.

A consideration of value-related issues properly includes: (i) how narrowly value should be defined;[96] (ii) whether value should be limited to *financial* value; (iii) if so, how broadly should *financial* be construed; (iv) how certain must value be (*e.g.*, how liquid; when should value be measured,[97] and should the potential value of consequential damages be included);[98] and (v) does the

---

93 "That which is yet to be executed or performed; that which remains to be carried into operation or effect; incomplete; depending upon a future performance or event." BLACK'S LAW DICTIONARY 680 (4th ed.1968).

94 Historically, cryptographic methods have (for other than national security purposes) largely only been required, or largely implemented, for financial purposes.

95 Some advocates of the broad view argue that even this standard is too weak. Instead, they propose that any transactions of "commercial significance" or some other more encompassing standards should be used.

96 Should the law focus on *clear value, face value, fair and equitable value, market value, true value,* or something else?

97 In an action to recover chattel, "value" means value at time of trial, not at time of seizure thereof. Spear v. Auto Dealers' Discount Corporation, 278 N.Y.S. 561 ( ).

98 Compare U.C.C. § 4A-305 ("Liability for Late or Improper Execution or Failure to Execute Payment Order") and U.C.C. § 2-715 ("Buyer's Incidental and Consequential Damages").

definition of *sensitive information* under the Computer Security Act of 1987 necessarily broaden the scope of value for such purposes?

A value limitation is ostensibly one of the most specific and well understood criteria. For example, statutes of frauds prescribe dollar limits, such as the $500 threshold of U.C.C. § 2-201. Another example is the Federal Acquisition Regulations ("FARs") which provide for a $25,000 threshold[99] and permit telephone bids/proposals or orders in an amount up to $2,500.[100] Federal money laundering regulations require reporting if a $10,000 daily aggregate amount is exceeded.[101] Specifying a baseline value has been criticized as both arbitrary and difficult to enforce; but, there are administrative rulings and interpretations that provide guidance and mitigate potential abuse of aggregate requirements.[102]

3. Costs of Implementation - Whether and how costs of security should impact baseline criteria are important issues to resolve. It is impossible to consider meaningfully the cost of resolving a problem until the nature of the problem and the underlying *requirements* are articulated. Premature consideration of costs may eliminate viable solutions; yet, intensive focus on cost (sometimes to the exclusion of all other factors) has been the linchpin for policy and legal reform efforts. The cost debate focuses on whether the use of cryptographic methods are a necessary component of "reasonable security procedures"[103] and whether the costs associated with cryptography are too burdensome to require.[104]

   This "crypto cost debate" has two main camps. Proponents of wide-spread cryptography usage argue that (i) only cryptography can adequately protect against the threats in open systems and ubiquitous computing environments, and (ii) because the costs of cryptography will decrease with increased usage, cryptography is a viable, indispensable, and appropriate requirement. Opponents of wide-spread cryptography usage argue that (i) conventional paper-based practices are fallible and consequently computer-based practices

---

99 48 C.F.R. § 13 (Small Purchase and other Simplified Purchase Procedures) (1992).

100 FAR 14.201-6(g)1 and 15.407(e)(1). The Defense FARs Supplement, 48 C.F.R. § 208.405-2 (allowing for oral procurement ordering from federal supply contractors).

101 31 C.F.R. § 103 (1990); 31 U.S.C. § 5315 (reports on foreign currency transactions).

102 *E.g.*, Administrative Rulings, Interpreting Treasury's Currency and Foreign Transactions Regulations, Fed. Reserve Reg. Serv. 88-1 (June 22, 1988).

103 Although the debate is focused on cryptography, a substantial proportion of fraud is traceable to inadequate conventional controls. Superior conventional controls would largely protect against such fraud (excluding the open systems issues). In this respect, the costs associated with implementing proper management controls may dwarf the costs of cryptography.

104 "When there is a homogeneous nationwide EFT network with standardized security techniques, it will become increasingly "cost effective" for criminal elements to develop the technology required to defraud the system, because this technology, once developed, could be applied nationwide against the cardholders of hundreds or even thousands of financial institutions." ANSI X9.9 Retail PIN Standard, § A.3. (Amer. Banker's Assn. 1982).

need not be any better,[105] and (ii) the costs of cryptography are greater than the costs associated with protecting conventional media.[106] Since this debate continues to obfuscate the rational development of policy and rules for computer-based media, cost issues deserve further examination.

Notwithstanding this debate, the commercial information security marketplace, and particularly the commercial cryptographic marketplace, are undergoing substantial changes which impact the accuracy of the cost analysis.[107] There is little rigorous publicly available analysis of the costs of implementing and using cryptographic methods.[108] A cost analysis for implementation of cryptography may include the additional costs, if any, incurred as a result of:

---

105 This argument may fail to account for the new and improved tools, as well as the possibilities offered by modern technologies. *See supra ABA* Resolution § (a) in Section II.b. of this paper, (encouraging appropriate legal and professional education).

106 Some proponents of the substantial use of cryptography retort by asking whether cryptography is more costly than a courier or a safe to protect an original?

107 Although market-based arguments against implementing new or stronger security mechanisms prevail, there is evidence that market demand for security products appears to have accelerated considerably. This position is cautiously, yet optimistically, presented in light of the many "false starts" which security market pundits' reports have historically missed.

108 For example, NIST plans to "[i]nvestigate the economic interests involved in the DSS." Miles Smid, "draft Response to comments on the NIST proposed digital signature standard," presented at Crypto '92 (Santa Barbara, Aug. 17, 1992) at 13. Note that the Data Encryption Standard (DES) "reflects hundreds of millions of dollars in investment," Geoffrey Turner, SRI, quoted in "Board to review U.S. policy on use of cryptography." Network World, Sept. 21, 1992 at 92.

SOURCE OF COST	APPLICABLE COST CONSIDERATIONS
•Crypto. software licensing •Certificate purchasing •Export filing process	•License negotiation •Certificate purchase costs •Legal and technical fees for export license •Perhaps these are diminishing issues if Software Publisher's Association-type policies & agreements proliferate, and export reform continues
•Additional cryptographic communications overhead	•Size of transactions (if transaction volume is great and the size of each such transaction is small proportionally, cost is a greater factor) •Communicating certificates/CRLs, etc. •Interoperable functional standards implementation
•Professional training, staffing and support[109]	•Comparatively few practitioners of the art •Considerable learning curve •Technical development nontrivial & highly variable •Problems in reaching agreement on implications of certificates •User training and servicing
•Additional processing[110] and storage	•CRL, certificate and message signing and verification •Host-based cycles (expensive compared to PCs) •Time sensitivity of subject data (a big factor)
•Key and certificate management and operation •Export "diversion in place" oversight[111]	•Liabilities of certificate issuer •Bonding and liabilities of "organizational notaries" •Issuance and revocation procedures, security and audit •Drafting and executing agreements and policies •Configuration management

TABLE 4 - SURVEY OF COSTS IN IMPLEMENTING CRYPTOGRAPHY

Another cost issue requiring resolution is whether governments will develop, or make agreements with providers to supply cryptographic software to small businesses or to the disadvantaged. If so, would such software distribution be viewed as illegally "in competition" with private enterprise. Recent events associated with "enhanced" or "value-added" information service provision by the Federal Maritime Commission and other agencies highlight this point.[112] Finally, differences between private and public policy objectives should be

---

109 One example is the training requirements under the Computer Security Act at 1987 Fed. Reg. 26,940 (June 12, 1991).

110 *See* Ronald L. Rivest, "On NIST's Proposed Digital Signature Standard," PROCEEDINGS OF THE SECOND CPSR CRYPTOGRAPHY AND PRIVACY CONFERENCE (Washington, DC, June 1, 1992) § 4.5.5 (providing an analysis of cryptographic processing costs and notes "an approximate doubling of computer power (per dollar) every two years, and an approximate increase of a factor of 4500 after twenty-five years . . . . In the year 2017, I expect computer power will be about 5000 times cheaper than it is now.").

111 An example of this is the costs associated with any requirements imposed on a network, and the costs to monitor or prevent actively the export of controlled technical data from the U.S. under the Export Administration Regulations.

112 *See generally* O.M.B. Management of Federal Information Resources Proposed Revision of OMB Circular A-130. 57 Fed. Reg. (No. 83) 18,296 (Apr. 29, 1992); Tariffs and Service Contracts, Federal Maritime Commission, 57 Fed. Reg. 36,268-36,311 (Aug. 12, 1992).

considered since conflicting agendas affect the choices available in designing model security baselines.[113]

4. Private vs. Public - Another consideration is whether, and how, baselines in the private and public sectors should vary. For example, a private sector "business risks" model may not be necessarily applicable to public sector obligations in which public servants play a non-profit and fiduciary role to the public at large. In such an environment, there may be a more compelling basis for strong security.

5. Present vs. Future - Where costs of computing continue to decrease rapidly, where availability of computers and security mechanisms continue to increase rapidly, and where there is growing confidence that "open systems" environments will become typical, should baseline requirements be skewed towards the present or the future (assuming that any requirements necessarily cannot be totally neutral as to their placement in time)?

6. Conflicting Security Requirements - Where baseline security requirements (such as statute, regulation or agreement) conflict with a particular transaction's special security requirement(s), the special requirements should preempt baseline requirements.

7. The Party(ies) Requiring Protection or Assurances - Whether the party requiring assurances or protection (*e.g.*, against revocation or repudiation) is either the originator, the recipient or a third-party beneficiary should be considered. For example, if the originator requires specific security assurances, security requirements can arguably be less stringent than where the recipient also requires assurances. This is because the originator is in the better position to control the type and extent of the security applied.[114] The recipient must either accept or reject that which the originator sends. TABLE 5 presents a simplified (perhaps over-simplified) comparison of various document types, the effects of which should be reflected by the Model Security Baseline.[115] TABLE 5 is necessarily subjective -- because the primary beneficiary of security will depend upon the particular circumstances.

---

113 For example, government goals in information security are typically not geared toward profit-oriented risk taking, but rather toward the prevention of fraud or other loss.

114 Originators may (depending on the implementation) optionally include cryptographically enhanced security (*e.g.*, digital signatures) for their own protection even where not legally required to do so. Whereas, in the absence of agreement or rule, the recipient is at the mercy of the originator.

115 *See infra* Section II.c. The Baseline adopts the term *message* for consistency with international standards. The term *transaction* or other descriptive term can be substituted by the user.

TYPE OF TRANSACTION	ORIGINATOR	RECIPIENT	BOTH	3RD PARTY BENEFICIARY
Complaint	X			
Credit EFT		X		
Debit EFT	X			
Deed / Will		X		X
Hazardous Waste Manifest	X			X
"I.O.U."		X		X
Notice			X	
P.O. / Contract			X	
Power of Attorney			X	X

TABLE 5 - PRIMARY BENEFICIARY OF SECURITY

c. A Model Security Baseline

The following Model Security Baseline ("Baseline") is presented as one approach that contributes to the development of rules affording greater certainty for the following risk assumptions. The Baseline assumes that the transactions are largely procurement or commercial in nature, and that the anticipated electronic commerce environment may include open systems. For simplicity, the Baseline creates three classes of messages:[116] *Level 1*, *Level 2*, and *Level 3*, each requiring incrementally stronger security, such as the use of cryptographic methods for authentication, integrity, and confidentiality purposes.[117] Three levels are within the boundaries of workable rule-making. Where more than three classes of security are desired or required, greater granularity in the levels, or additional levels with stronger or weaker characteristics can be developed responsively. The Baseline also contemplates greater specificity in subsequently derived rules using the Baseline as a tool.[118]

Both legal and computer security circles have expressed concern that security requirements should be separated from the specific security technologies implemented.[119] Although the following Baseline may be critiqued as providing nadequate separation, it provides comparatively general (and flexible) requirements.

---

116 These three classes of transactions are substantially consistent with the three classes of information in electronic form presented in TABLE 1 - COMPARISON OF SIGNED WRITINGS AND ELECTRONIC INFORMATION, *supra* Section II.c.

117 The Baseline is intended to help navigate through the pivotal decision (and perhaps the most difficult policy controversy) of whether or not to require the use of cryptographic-based security mechanisms. The Baseline is reprinted in Appendix 1 without footnotes and other distractions.

118 The Baseline provides a practical interface between policies and detailed rules. For example, the Baseline provides a roadmap for enforcing a security policy, and yet, it purposefully refrains from detailing cryptographic key size, levels of passwords, algorithms, whether hardware is needed to implement cryptography and other legal and security techniques, parameters and requirements. *See supra* TABLE 3 - RELATIVE LEVELS OF ABSTRACTION.

119 Early drafts of the ABA Resolution, *supra*, expressly considered cryptographic technologies. This consideration precipitated concern within the legal community that by mentioning cryptographic technologies (i) the failure to use them would create exposure, and (ii) the rules would become antiquated prematurely.

A MODEL SECURITY BASELINE - LEVEL 1

Section 1 - *Level 1 Message Attributes*. An electronically created, communicated, or stored message of the parties shall be considered a "Level 1" message where 1.a, 1.b, 1.c and 1.d are applicable – singularly or in any combination:

1.a. the message(s) does not contain highly sensitive information,[120] proprietary trade secrets, or other information requiring strong privacy protection;

1.b. the value[121] of the message(s) [over any [thirty (30)][122] day period] [as established by the parties] [is not expected to exceed] [does not exceed] [five thousand dollars ($5,000)][123] [twenty-five thousand dollars ($25,000)] [X dollars];

1.c. no applicable law specifies alternative security measures, or otherwise preempts the applicability of the Model Baseline for the specified message; or[124]

1.d. the message is not highly time sensitive.[125]

Section 2 - *Security/Reliability*. The security implemented for Level 1 messages shall include, at a minimum:

---

120 The use of the Computer Security Act of 1987 as a threshold for baseline criteria raises issues (and possibly problems) because most EDI information can reasonably be considered sensitive under the Act. The Baseline seeks to accommodate sensitive information under the act – providing incrementally stronger security in its Levels.

121 Value is intended to mean actual or fair market value. Notwithstanding this definition, legal damages, the value of a loss to society (*e.g.*, environmental pollution – potentially intangible or difficult to ascertain), as well as issues of consequential damages should also be considered. *See infra* and 98; Evra Corp. v. Swiss Bank Corp., 673 F.2d 951 (7th Cir. 1982) (failure of bank to properly handle telex wire transfer not liable for consequential damages because it had not been placed on notice of special circumstances giving rise to them).

122 The [bracketed] portions of text in the Baseline indicate their optional character. In fact, as a model, all provisions in the Baseline are ultimately optional.

123 The $5,000 is intended to be an aggregate amount. Its purpose is to prevent "splitting" large orders into multiple smaller ones. The use of a value limit on multiple transactions has proven difficult to enforce because the anticipated value/volume for a future time period is speculative. Inflation will render the $5,000 less important over time. A link to a government price index, such as the consumer price index might be useful. The author acknowledges that some knowledgeable legal and technical experts believe that an aggregate amount is either unnecessary or inappropriate.

124 Additional criteria could provide that: "the business situation does not present unusual elements which tend to increase the risk above normal levels." However, determining the parameters of "normal levels" could be difficult or fruitless.

125 *See supra* Section II.e. TRUSTED ENTITIES AND TIME STAMPING, regarding applications requires greater proof of timeliness. *E.g.*, in *Interactive EDI*, "[f]aster EDI is a primary requirement. This is not only a requirement on the underlying communications methods, but on all functional entities within and between the trading partners . . . . response times of seconds or fractions of a second, as opposed to minutes or hours, will generally be required." RECOMMENDATION TO UN/ECE/WP.4 ON INTERACTIVE EDI WITHIN THE CONTEXT OF UN/EDIFACT, TRADE/WP.4/R.842 (July 21, 1992) at 8.

2.a. noncryptographic identification and authentication [*e.g.*, password-user ID];[126]

2.b. recognized controls to ensure authenticity,[127] integrity, [confidentiality,] and availability;[128] and

2.c. audit trails.[129]

Section 3 - *Legal Effect.* For all legal purposes, all Level 1 messages which are communicated pursuant to Section 2 shall be presumed [conclusively?[130]] to be "in writing," "signed,"[131] authentic, and enforceable to no less an extent than if such messages had been undertaken using conventional (paper-based) mechanisms.[132]

Level 2 and 3 messages require stronger security. The following Baseline affords Level 2 messages greater security. Enhancement of Level 1 requirements is achieved through the addition of the use of cryptographic methods for MACs or digital signatures, and optionally for stronger confidentiality protection. Additions and deletions to Baseline Level 1 messages are noted accordingly.

---

126 "Noncryptographic identification and authentication" requires greater specificity such as by reference to National Institute of Standards and Technology (NIST) or other authoritative guidelines. Depending upon the implementation, security should minimally be of the "C2" level where the passwords are associated with an individual. Class C2: Controlled Access Protection makes "users individually accountable for their actions through login procedures, auditing of security-relevant events, and resource isolation." DEPARTMENT OF DEFENSE TRUSTED COMPUTER SYSTEM EVALUATION CRITERIA, DOD 5200.28-STD (Dec., 1985)(hereinafter "DoD Trusted Criteria") at 15.

127 Levels 1 and 2 of the Baseline do not accommodate full non-repudiation because of their lack of a trusted time stamp. *See supra* Section II.d. and II.e.. (concerning non-repudiation and trusted time stamps).

128 Such controls should be comparable to recognized and appropriate criteria, *e.g.*, in the nature of certain requirements included within the Class C2 Security Policy. *See* DoD Trusted Criteria *supra* at 15.

129 Each entity participating in a transaction (*e.g.*, each trading partner and all intermediaries) should be required to keep an audit trail. *See generally* A GUIDE TO UNDERSTANDING AUDIT IN TRUSTED SYSTEMS, National Computer Security Center, NCSC-TG-001 Version 1 (July 28, 1987); BELDEN MENKUS and ZELLA G. RUTHBERG, CONTROL OBJECTIVES, (EDP Audit Foundation, 1990).

130 Issues associated with conclusive presumptions are discussed in Section IV. *infra* BURDEN OF PROOF AND PRESUMPTIONS.

131 Where the message's originator intended the message to be signed and properly communicated, otherwise the presumption shall be that the transaction was intended to be in writing but not signed. A careful review of the purpose of each particular signature requirement must be undertaken; and the parties should be confident that the particular purpose of the signature requirement is met by the substituted electronic mechanisms.

132 *See* BAUM AND PERRITT, *supra* at 185-186.

A MODEL SECURITY BASELINE - LEVEL 2

Section 1 - *Level 2 Message Attributes*. An electronically created, communicated, or stored message(s) of the parties shall be considered a "Level 2" message where 1.a, 1.b, 1.c and 1.d are applicable -- singularly or in any combination:

1.a. the message ^ contains highly sensitive information, proprietary trade secrets, or other information requiring strong privacy protection;

1.b. the value of the message(s) [over any [thirty (30)] day period] [as established by the parties] [is ^ expected to exceed] [^ exceeds] [five thousand dollars ($5,000)] [twenty-five thousand dollars ($25,000)] [X dollars];

1.c. ^ applicable law specifies alternative security measures, or otherwise preempts the applicability of the Model Baseline for the specified message; or

1.d. the message is not highly time sensitive.

Section 2 - *Security/Reliability*. The security implemented for Level 2 messages shall include, at a minimum:

2.a. noncryptographic identification and authentication [*e.g.*, password-user ID];

2.b. recognized controls to ensure authenticity, integrity, [confidentiality,] and availability;

2.c. audit trails; and

2.d. [message authentication codes (MACs)[133]], [digital signatures] [and/or encryption for confidentiality].[134]

Section 3 - *Legal Effect*. For all legal purposes, all Level 2 Baseline messages which are communicated pursuant to Section 2 shall be presumed [conclusively?] to be "in writing," "signed," authentic and enforceable to no less an extent than if such messages had been undertaken using conventional (paper-based) mechanisms.

The following Level 3 messages have attributes which require "trusted third party" security services. Additions and deletions to Level 2 are noted. The satisfaction of other legal requirements, such as negotiability, will require alternative security services.[135]

---

133 This may involve using secret key techniques such as DES (*see* FIPS-PUB 46-1). *See* FIPS-PUB 113 on MACs.

134 This may be accomplished through the use of public key-based or conventional key-based key management and key exchange mechanism to transmit/create secret session keys for privacy of messages.

135 A trusted record keeper is anticipated to be necessary to accommodate computer-based negotiable documents. *See* BAUM AND PERRITT *supra* at § 5.11 "-Documentary Transfers," and § 11.9 "-Negotiability and Bills of Lading" (addressing trusted record keeping mechanisms for negotiable documents); TABLE 1 - COMPARISON OF SIGNED WRITINGS AND ELECTRONIC INFORMATION, *supra* at Section II.c.

A MODEL SECURITY BASELINE - LEVEL 3

Section 1 - *Level 3 Message Attributes.* An electronically created, communicated, or stored message(s) of the parties shall be considered a "Level 3" message where 1.a, 1.b, 1.c, 1.d and 1.e are applicable – singularly or in any combination:

1.a. the message ^ contains highly sensitive information, proprietary trade secrets, or other information requiring strong privacy protection;

1.b. the value of the message(s) [over any [thirty (30)] day period[[as established by the parties] [is ^ expected to exceed] [^ exceeds] [five thousand dollars ($5,000)] [twenty-five thousand dollars ($25,000)] [X dollars];

1.c. ^ applicable law specifies alternative security measures, or otherwise preempts the applicability of the Model Baseline for the specified message;

1.d. the message is ^ highly time sensitive; or

1.e. an acknowledgment by a notary public, or comparable "stronger" proofs or certifications is required.

Section 2 - *Security/Reliability.* The security implemented for Level 3 messages shall include, at a minimum:

2.a. noncryptographic identification and authentication [*e.g.*, password-user ID];

2.b. recognized controls to ensure authenticity, integrity, [confidentiality,] and availability;

2.c. audit trails;

2.d. [message authentication codes (MACs)], [digital signatures] [and/or encryption for confidentiality]; and

2.e. electronic notarization (time stamping and [MAC[136]] [digital signature]) by a trusted entity.

Section 3 - *Legal Effect.* For all legal purposes, all Level 3 Baseline messages which are communicated pursuant to Section 2 shall be presumed [conclusively?] to be "in writing," "signed," authentic and enforceable to no less an extent than if such messages had been undertaken using conventional (paper-based) mechanisms.

As presented, Section 3 - *Legal Effect* (of all three Baseline levels) focuses on assuring that computer-based messages are afforded comparable legal effect to paper-based messages. However, because Baseline Levels 2 and 3 use incrementally stronger security mechanisms (than in Level 1) that provide greater assurances of

136 There is not yet a viable infrastructure to support symmetric-based key management where several hundred thousand parties utilize a security mechanism. Also, notarization using MACing with symmetric key technology requires that verification of notarization must be provided exclusively by the notary since keys in such an implementation cannot be shared.

trustworthiness, there is a compelling basis for providing other beneficial legal effects within Section 3 - *Legal Effect*. Consequently, as an alternative, Baseline legal effects should provide incrementally stronger legal presumptions and burden allocations. For example, where a party used a digital signature, the authenticity and integrity of the computer-based information should be more difficult to attack legally (or rebut) than if weaker security had been applied to the message. The following two sections consider these issues in more detail and present such a proposal.

IV. BURDEN OF PROOF AND PRESUMPTIONS

There is no satisfactory test for allocating the burden of proof in . . . any given issue. [137]

Scant attention has been paid to burden of proof and presumption issues in electronic commerce. This is unfortunate since, after all, proof issues are at the heart of the meaningful resolution of disputes. Burden of proof and presumption issues have been approached largely without meaningful consideration of the *dynamic* proof sets[138] necessary to accommodate transaction-oriented environments. Dynamic proof sets differ sharply from the relatively *static* proof sets developed for record-oriented environments. While undeniably a daunting task, and an issue worthy of further study, burdens of proof and presumptions must be examined and integrated into a workable legal framework for electronic commerce.[139]

The development of electronic commerce rules are intimately affected by burden of proof requirements which consist of both the *risk of nonpersuasion* and the *duty of producing evidence.*[140] Burden of proof issues affect (i) electronic message reliability and genuineness, and (ii) admissibility and enforceability[141] of information in electronic form (*e.g.*, substituted for paper-based documentation).

In developing and evaluating rules governing electronic commerce, one must recognize that "[t]he burden of pleading [should be] allocated on the basis of pragmatic considerations of fairness, convenience, and policy, rather than on any general principle of pleading."[142] Yet, in many respects, the law's approach to the rules governing proof of facts at trial, as exemplified by the U.C.C., has been critiqued as:

---

137 GEOFFREY C. HAZARD, JR., CIVIL PROCEDURE, 322 (3rd ed. 1985) [hereinafter, "HAZARD"].

138 Telephone interview with Gregory P. Joseph, Esq., (Oct. 10, 1992).

139 For example, maritime law is rich in presumptions because there are often no witnesses to events on the high seas. Furthermore, cargo is, as a matter of course, passed through many hands internationally.

140 *See* HAZARD *supra* at 314. U.C.C. § 1-201(8) states that the *Burden of establishing* "a fact means the burden of persuading the triers of fact that the existence of the fact is more probable than its non-existence."

141 Electronic commerce legal commentators have often focused either on "enforceability" or on "evidentiary value."

142 HAZARD, *supra* at 323.

> remarkably casual, indeed almost haphazard. There are no general provisions constructing the evidentiary relationships of the parties, and the UCC's specific rules are insufficient to provide guidance on a host of significant and recurring problems. Predictably, the result has been that the goals of consistency and clarity in commercial law have not been achieved in the important area of evidentiary proof rules.[143]

One rule allocates the burden of proof to the party having the readier access to knowledge about the fact in question.[144] In electronic commerce, this party may vary considerably depending on the computer involved, communications architecture, applications, and the party intended to benefit from the electronic message, among other considerations.

The Federal Rules of Evidence delineate presumptions.[145] Presumptions are "occasionally used to refer to the logical inference of one fact from the existence of another."[146] For example, "[i]f Smith mails at a postbox a letter to Jones, with proper address and postage on the envelope, the trier may infer that Jones received the letter."[147] Similarly, "[i]t has been declared that there is a presumption, not conclusive, of prompt delivery of a letter mailed in the absence of evidence to the contrary."[148] "The degree of persuasion required is also sometimes manipulated as a handicap against disfavored contentions. Thus if a claim is presented that a written contract was orally modified, the party claiming the modification must in some jurisdictions prove its contention by clear and convincing evidence."[149]

"What, then, are the bases upon which courts or legislatures will create presumptions? For the most part they are the same kinds of reasons that influence the allocation of the production burden generally, and these may be summed up as reasons of convenience, fairness, and policy."[150] Additionally, distinctions in

---

143 Ronald J. Allen and Robert A. Hillman, *Evidentiary Problems in - and Solutions for - The Uniform Commercial Code*, 1984 Duke L. J. 92, 93.

144 HAZARD, *supra* at 324.

145 "In all civil actions and proceedings not otherwise provided for by Act of Congress or by these rules, a presumption imposes on the party against whom it is directed the burden of going forward with evidence to rebut or meet the presumption, but does not shift to such party the burden of proof in the sense of the risk of nonpersuasion, which remains throughout the trial upon the party on whom it was originally cast." FED. R. EVID. 301 "PRESUMPTIONS IN GENERAL IN CIVIL ACTIONS AND PROCEEDINGS."

146 9 WIGMORE *supra* at § 2492; *See* F.A.R. Liquidation Corp. v. Brownell, 140 F.Supp. 535 (D.DE 1956) (permitting inference based on fact established by direct or circumstantial evidence of time telegram communicated).

147 HAZARD, *supra* at 326.

148 Franklin Life Ins. Co. v. Brantley, 165 So. 834 (AL 1936); *see* Kiker v. Commissioner of Internal Revenue, 218 F.2d 389, 393 (4th Cir. 1955) (there was no presumption that a letter was delivered in the ordinary course of the mails where address was not proper).

149 HAZARD, *supra* at 325.

150 HAZARD, *id.* at 328.

constitutional and procedural requirements for burdens of proof and presumptions in civil versus criminal proceedings must be considered.[151]

The use of presumptions affecting validity or enforceability of information in electronic form are widespread in EDI agreements. One example is the Model Electronic Payments Agreement and Commentary ("Model Agreement"), which states that "[t]he receipt by the sender of an acknowledgment from the recipient shall constitute *conclusive evidence* that the subject communication was received and is syntactically correct."[152] The practical effect of a conclusive presumption[153] is to excuse the sender from proving receipt where the proof is entirely in the recipient's control, perhaps to do otherwise would render EDI commercially ineffective.[154]

To what extent should *conclusive presumptions* be subject to attack?[155] How much evidence should be required to disprove or shift a presumption? By analogy, take the case of the mails. "If, for example, the addressee of a properly mailed letter testifies that he or she never received it, that testimony would, if believed, justify a finding of nonreceipt." "[T]he destruction of the presumption would not, however, compel a finding of non-receipt because a properly addressed letter is so likely to reach its destination that a rational inference may be drawn that it did so."[156] Should such a presumption hold in electronic commerce matters? And, to what extent should or must there be a *rational connection* between the fact presumed and the fact proved? To illustrate the approaches taken in many domestic and international model electronic commerce agreements, the Commentary to the Model Electronic Payments Agreement presents the following addition presumptions:

> Validity and Enforceability. Neither party shall contest the validity or enforceability of Transaction Sets or notices communicated pursuant to this Agreement on grounds related to the absence of paper-based writings, signings or originals.

---

151 *E.g.*, Hazard notes "an intermediate test which is occasionally applied in civil controversies" – "clear and convincing evidence." *See* Notes of Advisory Committee of the 1972 Proposed Rules, FED. R. CIV. P. 301.

152 MODEL AGREEMENT, *supra* at Section 6.3, and MODEL EDI TRADING PARTNER AGREEMENT *supra* at § 2.2., Comment 7.

153 "The *conclusive presumption* is not really a procedural device at all. Rather it is a process of concealing by fiction a change in the substantive law. When the law conclusively presumes the presence of B from A, this means that the substantive law no longer requires the existence of B in cases where A is present, although it hesitates as yet to say so forthrightly. (emphasis added). 9 WIGMORE *supra* at § 2492; and Gordon and Tenenbaum, "Conclusive Presumption Analysis: The Principal of Individual Opportunity," 71 NW. U. L. REV. 579 (1976).

154 However, the parole evidence rule does permit the voluntary adoption of a "super parole evidence rule" that prevents the parties from using evidence of future oral modifications. *See* U.C.C. § 2-202 "Final Written Expression: Parol or Extrinsic Evidence."

155 By definition, *conclusive presumptions* are irrefutable, yet in practice, they are sometimes refutable.

156 HAZARD. at 330; 9 WIGMORE, *supra* at § 2489. Given the various documented instances where mail is destroyed or delayed, this presumption is suspect.

Each Transaction Set and notice communicated in electronic form pursuant to this Agreement shall be considered to be:

(a) "in writing" and "written" to an extent no less than if in paper form;

(b) "signed" where the signer includes data intended as a signature [as agreed among the parties] to an extent no less than if conventionally undertaken with pen and paper; and

(c) an original.[157]

Examples of other instructive presumptions include the following:

i. If EDI messages are transmitted in accordance with an authentication procedure such as a digital signature, they shall have, between parties, a comparable evidentiary value to that accorded to a signed written document.[158]

ii. In an action with respect to an instrument, the authenticity of, and authority to make, each signature on the instrument is admitted unless specifically denied in the pleading. If the validity of a signature is denied in the pleadings, the burden of establishing validity is on the person claiming validity, but the signature is presumed to be authentic and authorized unless the action is to enforce the liability of the purported signer and the signer is dead or incompetent at the time of the trial of the issue of validity of the signature.[159]

iii. If there is a discrepancy between the terms of the payment order transmitted to the system and the terms of the payment order transmitted by the system to the bank, the terms of the payment order of the sender are those transmitted by the system.[160]

iv. A document in due form purporting to be a bill of lading . . . or any other document authorized or required by the contract to be issued by a third party shall be prima facie evidence of its own authenticity and genuineness and of the facts stated in the document by the third party.[161]

---

157 MODEL AGREEMENT, *supra* at § 6, Comment 13.

158 TEDIS, EUROPEAN MODEL EDI AGREEMENT, ART. 10 (Final Draft, 1991).

159 U.C.C. § 3-308(a) ("Proof of Signatures and Status as Holder in Due Course.") "The presumption rests upon the fact that in ordinary experience forged or unauthorized signatures are very uncommon, and normally any evidence is within the control of, or more accessible to, the defendant." *Id.* Official Comment 1.

160 U.C.C. § 4A-206 ("Transmission of Payment Order through Funds-Transfer or Other Communications System.").

161 U.C.C. § 1-202 ("Prima Facie Evidence by Third Party Documents.").

The Model Security Baseline[162] includes presumptions which may vary, and which deserve further scrutiny. One immediate issue is whether the Baseline (as well as the various EDI-related model agreements) should delve further into burdens of proof and other evidentiary matters. If incrementally greater security mechanisms are used (such as in Model Baseline Levels 2 and 3), *why should not the parties receive incrementally increased presumptions as to the admissibility, credibility and weight to be afforded such messages?*[163] For example, TABLE 6 proposes replacing (alternatively, adding to) the Model Baseline's Section 3 - *Legal Effect* with the following presumptions for certain classes of messages. This is intended to provide a more dynamic risk-based model, and provide stronger security users with appropriate and commensurate benefits.[164]

MODEL BASELINE SECTION 3, LEVEL:	PRESUMPTION	(SUBSTITUTE SECTION 3 - *LEGAL EFFECT*)
1	Rebuttable Presumption A	Shifts burden of proof to rebut presumption by *a preponderance of the evidence*
2	Rebuttable Presumption B	Requires *clear and convincing* proof to rebut presumption of authenticity
2A (alternative to 2)	Rebuttable Presumption C	Requires proof *beyond a reasonable doubt* to rebut presumption of authenticity
3	Irrebutable Presumption	Presumption is conclusive regardless of the opponent's evidence

TABLE 6 - SUBSTITUTE MODEL BASELINE SECTION 3 - LEGAL EFFECT[165]

TABLE 6 may be preferable to the Baseline's *Section 3 - Legal Effect,* because the TABLE 6 presumptions are not inherently tied to conventional paper-based technologies. Finally, the increasing strengths of the presumptions in Table 6 are more dynamic than those of the Baseline and can be used in a multidimensional scheme.[166] Consequently, TABLE 6 deserves further consideration.

---

162 Section III.c., *supra.*

163 There is strong basis in the law for providing greater legal effect to documents which have been more strongly authenticated or secured; this is the case with self-proving wills and some statutes of limitations.

164 It has been comically suggested that "as you move much beyond three to four levels of burdens of proof, no one except Judge Wapner could possibly understand and effectively use it." Interview with Alfred I. Maleson, Prof. Emeritus, Suffolk Univ. Law School, in Boston (Nov. 4, 1992).

165 Transactions which do not satisfy the security criteria of Baseline Level 1 could, depending upon the legal scheme, be viewed as representing *simple presumptions* which shift the burden of going forward with the evidence, but do not change the burden of proof.

166 For example, a scheme could be developed where a Baseline Level 1 transaction uses Level 2 security and therefore responds to a stronger presumption.

V. INTEGRATING FORMALISTIC & EVIDENTIARY REQUIREMENTS

Legal requirements for information in electronic form are typically evaluated from one of two perspectives: (i) formalistic-related requirements (*e.g.*, focusing on requirements for, or the sufficiency of, substitutes for "signed writings"),[167] or (ii) evidentiary-related requirements (focusing on admissibility, credibility and proof issues).[168] Where these perspectives are either viewed in a vacuum or adopted without contemplating their interrelationship, the resulting perspective and rules are destined to be dysfunctional. Insufficient attention has been directed toward utilizing an integrated *cradle-to-grave* analysis of the total electronic commerce environment and its requirements. To aid such an analysis, FIGURE 1 presents a representative cradle-to-grave analysis of a transaction. FIGURE 1 segments electronic commerce transactions into four phases of legal import: Phase 1-Creation (includes processing), Phase 2-Communication, Phase 3-Verification (includes retention functions)[169] and Phase 4-Dispute Resolution [170]

---

167 *See* Section II.a. "Treatment in the Law" *supra.*

168 *See* Section IV. "BURDEN OF PROOF AND PRESUMPTIONS," *supra* .

169 Transaction record storage would logically follow Phase 1 - verification -- and verification might be undertaken following each use of the stored information.

170 In this hypothetical transaction: **[Phase 1]** a user creates information in electronic form to which some signature or authentication mechanism is used to satisfy legal requirements and to mitigate security threats. Then, optionally, the document is witnesses or cosigned or both, and if necessary, notarized (perhaps via a trusted crypto. box). **[Phase 2]** Next, the document is communicated to the intended recipient via third party service provider. The recipient then accesses and obtains the message. **[Phase 3]** The recipient then verifies the message for assurances of authenticity using one or more of a variety of verification techniques. Following verification, the recipient optionally can communicate an acknowledgment back to the originator such as a functional acknowledgment to notify the originator that the message was received and syntactically correct. Also, where the transaction is contractual in nature, the recipient can communicate an acceptance. **[Phase 4]** Should a dispute ensue, the parties present admissible evidence to the dispute resolution mechanism and seek to persuade the fact finder, in part, by the weight and credibility of the evidence. A decision by the fact finder completes the hypothetical.

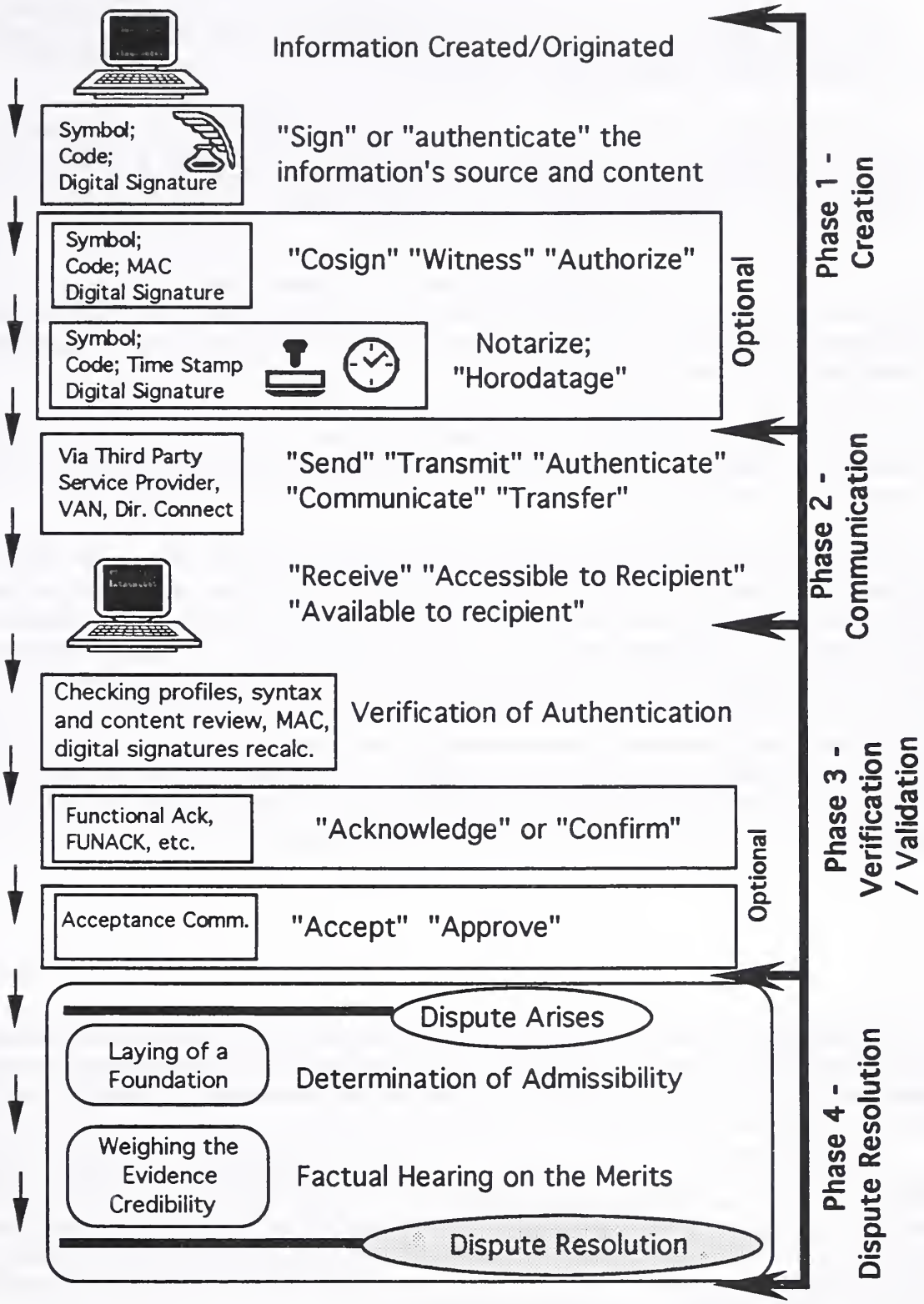


FIGURE 1 - A HYPOTHETICAL CRADLE-TO-GRAVE TRANSACTION

The remainder of this Section begins considering the following questions -- questions that deserve study beyond this paper:

a. If formalistic requirements are reduced or eliminated[171] (*e.g.*, at Phase 1, FIGURE 1), will the evidentiary requirements of laying a foundation (preliminary evaluation of authenticity and relevancy) necessarily shift to a factual determination of weight and credibility?[172]

b. If so, will such a shift either increase or decrease the total quantum of proof required (*e.g.*, at Phase 4, FIGURE 1) from either party. Further, will it qualitatively shift the *status quo* to the unintended or unjustified disadvantage of one of the parties?

c. Should the evidentiary requirements for laying a foundation be minimized, thereby further rendering the litigation to one of credibility and weight of the evidence?

If both formalistic and evidentiary foundational requirements are minimized, it is likely that a new risk will be created because the total required quantum of proof (weight and credibility) may increase. The party seeking to introduce a document cannot pregauge the extent of the required proof. The party cannot therefore rely on otherwise existing relatively static proof requirements. Absent definable and widely recognized formalistic requirements for electronic commerce, the current formalistic requirements for paper-based documents become less predictable. Theoretically, the potential evidentiary requirements, including the burden of proving transactions, become infinite. TABLE 7, presents some of the proffered relationships between formalistic, evidentiary foundational, and proof requirements.[173]

---

171 For example, these include requirements for a signatures, or their electronic analogs for the creation of enforceable documents in electronic form. If requirements for a signature are replaced by requirements for an electronic analog, then, the formalistic requirements remain, however, they simply take on a new form -- an electronic form.

172 The elimination of formalistic requirements is not out of step with modern legal developments. "What is valued is not form for form's sake, but useful form." LAWRENCE M. FRIEDMAN, A HISTORY OF AMERICAN LAW 278 (Touchstone Book, 2nd ed. 1965) "The statute of frauds survived; other formalities, which had no useful place, disappeared from the law of contract." "In general, sentiment and tradition had little place in commercial law; what survived was the fit and the functional." *Id.* at 279. It is precisely the signature, of course, which is alleged by many contemporary scholars and practitioners to be formalistic, unfit and dysfunctional.

Similarly, "[t]he advantage of the writing was not only that it furnished better proof. . . but also that it made it possible to enforce obligations for which there would otherwise have been no proof at all." OLIVER WENDELL HOLMES, THE COMMON LAW 262 (1881).

173 Perhaps the most tenuous of relationships is between foundational and proof requirements. An accurate description of the relationship is difficult to draft. However, the relative effects between formalistic requirements and evidentiary requirements are better substantiated.

Statute of Frauds or Comparable Formalistic Requirements		Relative Strictness of Requirements of Laying a Foundation for Admissibility		Anticipated Effect on the Quantum of Proof (Weight & Credibility) to Ensure Enforceability
Yes	+	Greater	=	Lesser
No	+	Greater	=	Medium
No	+	Lesser[174]	=	Greater

TABLE 7 - EFFECT OF DIFFERING FORMALISTIC & FOUNDATIONAL REQUIREMENTS

Some commentators propose that all information in electronic form should be admitted into evidence.[175] Under this view, the judicial process almost exclusively involves the fact finder determining the credibility of the evidence *without* the prerequisite of meaningfully laying a foundation. Alternatively, if a foundation were required, then it would be, a largely perfunctory requirement to minimize clearly irrelevant and prejudicial materials under Fed. R. Evid. 104(a) "Questions of Admissibility Generally" and Fed. R. Evid. 403 "Exclusion of Relevant Evidence on Grounds of Prejudice, Confusion, or Waste of Time", respectively. Consequently, the inquiry into reliability and trustworthiness would no longer be bifurcated into (i) laying a foundation as a prerequisite to admissibility, and (ii) determining the weight and credibility of the evidence.[176] Table 7 illustrates the anticipated dynamics of the trade-offs between these differing policies. One interpretation suggests that the diminution of formalistic and foundational requirements (the elimination of Statute of Frauds-like requirements and the relaxation of evidentiary foundation requirements) may not necessarily reduce electronic commerce barriers and costs. One euphemism which characterizes this concept in that there is *no free lunch*. What one tends to gain in the "Creation Phase" of the transactions (*see* FIGURE 1), is later lost by a commensurate increase in "Proof Phase" requirements.

VI. CONCLUSION

This paper recognizes the contribution of appropriate security techniques, procedures and practices to the legal efficacy of electronic messages and records. There is an inherent linkage between security and legal efficacy that is not adequately appreciated. The security of electronic messages and records is not only a business

---

174 The author recognizes that judiciary is likely to always demand some evidentiary foundation oversight.

175 One commentator advocates that "for business records virtually everything should be admissible, unless it is inherently unreliable - and even then I have doubts about the wisdom of creating a rule applying to EDI that would exclude any evidence . . . . Exclusion due to inadmissibility is a drastic sanction that can deprive a party of its fundamental proofs." Letter from George F. Chandler, III, Esq. to Michael S. Baum (Sept. 10, 1992) (on file with author).

176 In practice, however, the bifurcation has sometimes been blurred. "Any evidentiary shortcoming [in developing a foundation for admission of printouts from a computer retrieval system in drug prosecution] became a matter of weight to be given to the evidence rather than one of admissibility." U.S. v. Scholle, 553 F.2d 1109, 1124-25 (8th Cir.), *cert. denied*, 434 U.S. 940 (1977).

requirement,[177] but also is an underlying legal requirement. Defining this linkage is indispensable to the rational and pragmatic development of reliable electronic commerce. When the law determines what is sufficiently secure, it must consider the particular message's risks and purpose(s). Legal requirements should clarify *reasonable security procedures* without sacrificing needed flexibility. It is not a question of "having security" or "not having security" rather, it is a question of the *strength* of the security mechanisms implemented. When this legal-security linkage becomes broadly recognized, then the progress in the law which the electronic commerce community deserves and demands will begin.

***

177 "Clearly, security is an essential business requirement and is, therefore, at the heart of UN/EDIFACT." UN/EDIFACT Security JWG, Draft Rec. for Security (Jul. 1992).

APPENDIX - THE MODEL SECURITY BASELINE GRAPHICS

A SECURITY BASELINE - LEVEL 1

Section 1 - *Level 1 Message Attributes.* An electronically created, communicated, or stored message of the parties shall be considered a "Level 1" message where 1.a, 1.b, 1.c and 1.d are applicable -- singularly or in any combination:

1.a. the message(s) does not contain highly sensitive information, proprietary trade secrets, or other information requiring strong privacy protection;

1.b. the value of the message(s) [over any [thirty (30)] day period] [as established by the parties] [is not expected to exceed] [does not exceed] [five thousand dollars ($5,000)] [twenty-five thousand dollars ($25,000)] [X dollars];

1.c. no applicable law specifies alternative security measures, or otherwise preempts the applicability of the Model Baseline for the specified message; or

1.d. the message is not highly time sensitive.

Section 2 - *Security/Reliability.* The security implemented for Level 1 messages shall include, at a minimum:

2.a. noncryptographic identification and authentication [*e.g.*, password-user ID];

2.b. recognized controls to ensure authenticity, integrity, [confidentiality,] and availability; and

2.c. audit trails.

Section 3 - *Legal Effect.* For all legal purposes, all Level 1 messages which are communicated pursuant to Section 2 shall be presumed [conclusively?] to be "in writing," "signed," authentic, and enforceable to no less an extent than if such messages had been undertaken using conventional (paper-based) mechanisms.

A SECURITY BASELINE - LEVEL 2

Section 1 - *Level 2 Message Attributes.* An electronically created, communicated, or stored message(s) of the parties shall be considered a "Level 2" message where 1.a, 1.b, 1.c and 1.d are applicable -- singularly or in any combination:

1.a. the message contains highly sensitive information, proprietary trade secrets, or other information requiring strong privacy protection;

1.b. the value of the message(s) [over any [thirty (30)] day period] [as established by the parties] [is expected to exceed] [exceeds] [five thousand dollars ($5,000)] [twenty-five thousand dollars ($25,000)] [X dollars];

1.c. applicable law specifies alternative security measures, or otherwise preempts the applicability of the Model Baseline for the specified message; or

1.d. the message is not highly time sensitive.

Section 2 - *Security/Reliability.* The security implemented for Level 2 messages shall include, at a minimum:

2.a. noncryptographic identification and authentication [*e.g.*, password-user ID];

2.b. recognized controls to ensure authenticity, integrity, [confidentiality,] and availability;

2.c. audit trails; and

2.d. [message authentication codes (MACs), [digital signatures] [and/or encryption for confidentiality].

Section 3 - *Legal Effect.* For all legal purposes, all Level 2 Baseline messages which are communicated pursuant to Section 2 shall be presumed [conclusively?] to be "in writing," "signed," authentic and enforceable to no less an extent than if such messages had been undertaken using conventional (paper-based) mechanisms.

A SECURITY BASELINE - LEVEL 3

Section 1 - *Level 3 Message Attributes.* An electronically created, communicated, or stored message(s) of the parties shall be considered a "Level 3" message where 1.a, 1.b, 1.c, 1.d and 1.e are applicable -- singularly or in any combination:

1.a. the message contains highly sensitive information, proprietary trade secrets, or other information requiring strong privacy protection;

1.b. the value of the message(s) [over any [thirty (30)] day period[[as established by the parties] [is expected to exceed] [exceeds] [five thousand dollars ($5,000)] [twenty-five thousand dollars ($25,000)] [X dollars];

1.c. applicable law specifies alternative security measures, or otherwise preempts the applicability of the Model Baseline for the specified message;

1.d. the message is highly time sensitive, or

1.e. an acknowledgment by a notary public, or comparable "stronger" proofs or certifications is required.

Section 2 - *Security/Reliability.* The security implemented for Level 3 messages shall include, at a minimum:

2.a. noncryptographic identification and authentication [*e.g.*, password-user ID];

2.b. recognized controls to ensure authenticity, integrity, [confidentiality,] and availability;

2.c. audit trails;

2.d. [message authentication codes (MACs)], [digital signatures] [and/or encryption for confidentiality]; and

2.e. electronic notarization (time stamping and [MAC] [digital signature]) by a trusted entity.

Section 3 - *Legal Effect.* For all legal purposes, all Level 3 Baseline messages which are communicated pursuant to Section 2 shall be presumed [conclusively?] to be "in writing," "signed," authentic and enforceable to no less an extent than if such messages had been undertaken using conventional (paper-based) mechanisms.

Balanced Electronic Data Interchange Security

Irvin Chmielewski
Electronic Data Systems

Electronic Data Interchange is not a new technology. Industries have been using EDI in one form or another since the Seventies. The automotive industry is one of the areas that has been doing EDI for over twenty years and has transmitted hundreds of millions of transactions. The questions that are being asked today by new EDI implementors concerning appropriate levels of EDI security have been asked in the past. The answer that has emerged is the idea that a balanced approach is necessary for an effective security program that is cost effective to implement.

Has a balanced approach worked? The volume of Electronic Data Interchange transmissions being sent and received by the U.S. auto industry is not trivial; one major manufacturer, for example, currently sends and receives close to 1.5 million transmissions monthly. In discussions I've had, no one remembers a problem with a transmission being sent to the wrong location or any other problems due to unauthorized interception or tampering with transmissions.

Putting the highest level of security hardware and software into the processes is not always the best solution to a secure EDI implementation. A balanced and consistent use of various techniques must be achieved to insure secure transmissions, along with integrity and confidentiality of application data bases, while still encouraging the use of EDI solutions. An Information Security Strategy providing a set of technical rules, needs to be developed that ensures that applications will be secured in a consistent manner. By following a consistent but flexible approach security can be accomplished across dissimilar networked applications while reducing exposure for all connected applications and networks.

A recent development in systems architecture has been the design of client/server applications. The client/server application model benefits customers in areas such as flexibility, user-friendliness, specialization, economics and interconnection. Security of client/server applications also

requires a complete understanding of customer requirements as well as security standards. This understanding is based on risk analysis performed by the customer.

A key driver in the industry's EDI programs has been the goal of wide spread, almost universal usage. To ensure integrity, availability and confidentiality, applications both client/server and host-based, at the manufacturers and their trading partners, must provide full security of all information and resources. Not only Management Information Systems suppliers, but business customers must define, implement and support security in their applications.

EDI protection has been accomplished in most situations without any significant expenditure on high-tech security methodologies such as encryption. Yet, security of the overall process is of concern and is constantly being addressed. Analysis of the amount of risk involved in specific types of transaction set usage has lead to the use of sophisticated security methods of encryption and authentication in areas such as financial payments. A balance of high and low tech solutions to achieve EDI security has made the process workable in the real world where large numbers of trading partners are involved.

Just as the auto industry has learned that there is no single answer to the level of security, we have also learned that there are a number of places to apply security. Security of electronic data needs to be addressed in two main areas of the process. The first is data that is inside a trading partners' computer. The second is data security for information in the communications network.

Within application systems that will use EDI, the rules for authorized access and modification are the same as in any other sensitive application. Data security is essential. EDI capable systems must also maintain information concerning trading partners in addition to normal data. Trading partner information is used to set up and route the information transmissions. Under older systems, addresses were affixed to envelopes and sent by U.S. mail. Under an EDI process, basically the same thing is occurring; but, addressing may now

include passwords and mailbox identifiers needed to enter a trading partner's communication network.

An application's source for communication information, including computer address for a trading partner's mail box, network password and other delivery information, must be kept under the same rigorous security as the most sensitive data that will be sent to a trading partner. Why? Let us look at an example of two systems, one for distribution of a newsletter and the other for sending Purchase Orders, both by way of EDI. At first glance, the newsletter system would not require any security while the PO system would. Even if the trading partner and network information data were in files exclusive to each system, if an unauthorized copy was made of the newsletter addressing file, the addressing information that was on the PO addressing file would be compromised. Most companies have only one mailbox and password on a network. Larger companies may have multiple mail boxes, but usually due to geographical differences and not application differences. The trading partner and network information maintained for the low risk newsletter system can provide access to the PO system and therefore should be subject to the same controls as the PO system.

The keys to maintaining security within the value added network are shared by both the sender, the receiver, and the network. The value added networks, both private and proprietary; use logon id's and passwords for accessing the individual mail boxes. Most networks have requirements for a password to be changed over some period of time. It is this individually set password/mail box identifier combination that makes interception of a file difficult. Senders and receivers must maintain security over the passwords they use to access a network. Network providers must make the ability to change passwords easy to implement, and trading partners must change their own passwords on a frequent basis.

What if someone does get into a mailbox and intercepts a message? Mailboxes, when read, are emptied. Receivers know the processing schedules of their trading partners and usually go after their information in a timely manner. If expected information is not present, the receiver of the transaction set must immediately start resolving the problem. In situations,

where a permanent time schedule of transmissions has not been established, the use of an interchange acknowledgment is encouraged. For control purposes, use application acknowledgments to notify the sender that a transmission has been received and to echo back some key control total information. The acknowledgment is usually generated off trading partner tables not directly from received transactions. If a trading partner receives an acknowledgment for an unknown transmission, people can start making personal contact to understand what has occurred.

A certain amount of trading partner cooperation aids in the overall security of EDI activity. Applications must be thoroughly tested to produce accurate transmission addressing. Applications must be secured so that electronic addresses and other trading partner information can not be change by unauthorized methods. Passwords for entry into mailboxes must be kept confidential and should be changed on a frequent basis. Third party networks must be able to provide control over mailboxes and the files they contain to prevent access by unauthorized parties.

The Need for Risk Analysis

Robert V. Jacobson
President
International Security Technology, Inc.
New York City (212) 557-0900

Background

The purpose of Electronic Data Interchange (EDI) is to reduce or eliminate many common paper business documents...engineering specifications and drawings, requests for proposals, proposals, purchase orders, invoices and payment orders...and to permit business partners to exchange such documents electronically, computer-to-computer, with a minimum of human interaction. The objective is to reduce handling costs, the time to complete transactions, and processing errors. To be successful, EDI requires (1) a common set of specifications for structuring and processing these messages, (2) a large number of private and public organizations that have agreed to use EDI as the primary means of information interchange, and (3) EDI processing systems and procedures that ensure that the reliability and integrity are maintained at an adequately high level. As these goals are being met, EDI is becoming increasingly important in the conduct of business, particularly for large businesses and their government agency customers. For example, the National Electronic Information Corp., a consortium of health care payers, recently awarded a contract to implement a real-time

information network for medical claims processing with the goal of saving more than $4-billion per year.
As the use of EDI becomes widespread, risks will inevitably increase in three basic operating areas: service interruptions, unauthorized disclosure, and fraud. The more successful EDI becomes, the greater is the disruption when EDI service interruptions occur. As EDI processing systems grow in size, it becomes increasingly more costly to provide timely recovery provisions for them, and manual fall back schemes become impossible. The greater the volume of payment transactions, the more attractive EDI systems become to embezzlers. As the use of EDI decreases human participation in transaction processing, it becomes more difficult to detect frauds quickly. At the same time the sheer volume of transactions makes penny shaving frauds profitable for embezzlers. Experience with on-line credit reporting systems has shown that there are significant rewards to persons who can access to personal data on a regular basis for improper purposes. In short, risk exposures will grow as the use of EDI grows.

Why Risk Management Is Important To EDI Systems

If all EDI systems were the same: the same size, transaction volume, information sensitivity, urgency, monetary activity level, and operating environment, it would be possible to define an appropriate security program and apply it to all EDI systems without further consideration. However, this is not the case. EDI systems vary in all the dimensions just enumerated. Consequently, it is not possible to define a single security program for all EDI systems. EDI risks can only be managed effectively with rational risk management. Perfect security (nothing will ever go wrong) is infinitely expensive, and so it cannot be a rational design goal. On the other hand, inadequate security leads to unnecessary risk-related losses.

Risk management has two basic objectives:

1) Optimize the selection, and implementation of security measures based on a rational assessment of risks. "Optimize" in this context means that the objective is to implement security measures so as to minimize the sum of future risk losses and security expenditures.

2) Protect against catastrophic losses. A catastrophic

loss for a private sector firm would be a loss in excess of its equity. In other words, if the event occurs, however unlikely the occurrence may be, the loss will bankrupt the firm. While the concept of bankruptcy does not apply to government agencies, such agencies have a responsibility to the taxpayers to mitigate exposures to material losses.

To meet these two risk management objectives it is essential to evaluate the risks to which an EDI system is exposed in order to measure the utility of proposed security measures, and to identify potentially catastrophic risks. This process is commonly referred to as a risk assessment. A risk assessment uses three kinds of input data:

1) The rate of occurrence of the threats to the system being analyzed.

2) The potential for loss (loss potential) associated with each of the functions performed by the system, and each of the assets controlled by the system.

3) The vulnerability of the functions and assets to each of the threats that have an impact on them. (Note that a vulnerability by itself is not significant. While an asset may, indeed, be vulnerable to a given threat, the vulnerability is not significant unless the threat is expect to occur. This is why a vulnerability assessment so-called may yield useful insights about the state of existing security, but it is NOT the equivalent of a risk assessment.)

In the real world, details of threats, functions, and assets can be quite complex. Consequently, a key part of the risk assessment process is to construct a model of the EDI system being analyzed that aggregates the threats, functions and assets in to manageable groups.

The Need for Quantitative Risk Assessment.

The cost of security measures is stated in monetary terms. Therefore, one must state the benefit of security measures (the expected reduction in future losses) in monetary terms in order to compare cost and benefit. This is the basic reason for performing quantitative risk assessments. Installing a security measure is not prudent unless its

benefit outweighs its cost. The benefit of a security measure is the effect it will have on future losses. The purpose of a Quantitative Risk Assessment (QRA) is to generate an estimation of the losses that will occur in the future using quantitative estimates of the threat occurrence rates, loss potentials, and vulnerabilities as defined by the model of the system. Losses are expressed in two ways:

1) Annualized Loss Expectancy (ALE). ALE is the estimated loss expressed in monetary terms at an annual rate, for example, dollars per year. The ALE for a given threat with respect to a given function or asset is equal to the product of the estimates of occurrence rate, loss potential, and vulnerability rate. If the threat's occurrence rate is less than once per year, the ALE must be understood to represent the relative significance of a threat with respect to other threats. For example, imagine that the occurrence rate of a threat is estimated to be once in ten years, and its ALE is estimated to be $1,000 per year. This does not mean that the threat will cause a $1,000 loss in each of the next ten years. Instead, it will cause a $10,000 loss in one of the next ten years, but the specific year of occurrence cannot be determined.

   However, if one estimates ALEs for two threats as $1,000 per year and $100,000 per year respectively, all other things being equal, the second threat is clearly far more significant than the first one. Thus, ALE is a useful tool for ranking risks, even though confidence in ALE estimates tends to decrease as occurrence rate decreases. Even when quantitative estimates are relatively uncertain, they provide more risk management guidance than purely qualitative estimates of risk.

2) Single Occurrence Loss (SOL). SOL is the loss expected to result from a single occurrence of a threat. It is calculated for a given threat by summing the products of all the loss potentials of a system and their vulnerabilities with respect to the threat. Since SOL does not depend on an estimate of the threat's occurrence rate, it is particularly useful for evaluating rare but damaging threats.

In short, ALE is useful for addressing relatively frequent threats, and SOL is used to evaluate rare threats.

QRAs are used in three ways:

1) By comparing an as-is ALE with an ALE that assumes the presence of one or more proposed security measures, one can estimate the payback of the security measures. Obviously, the greater the ratio of the payback (reduction in ALE) to the cost of a security measure, the more valuable it will be.

2) Estimates of the losses that result from a single occurrence of a threat, its SOL, can be used to identify the potentially fatal threats as mentioned above.

3) ALE can be used to rank functions and assets relative to one another, and to rank the threats relative to one another, so as to prioritize plans to for asset protection, disaster recovery, and business resumption planning.

Obstacles To The Use of Quantitative Risk Assessments.

OMB Circular A-130, Appendix 3 specifically requires that risk assessments be performed for data processing systems. While Government agencies have to some extent performed risk assessments, they have made little use of risk assessments to manage the risks of their data processing systems. One can postulate several reasons:

1) Perhaps the most important reason is the reluctance to invest resources in the reduction of future losses. While regulations like A-130 include security requirements, the budgeting process typically does not include the explicit evaluation of risk, and the identification of appropriate security programs. When security resources are limited, they will be applied to the most obvious risks, and analysis of risks is assumed to be superfluous.

2) Government agencies do not have balance sheets, and government officials don't emphasize return on investment as much as the private sector. Likewise asset protection is not a major management goal. For example, private sector managers regularly balance outside insurance protection against self-insurance whereas the Government automatically self-insures.

3) Managers of Government data processing systems, being for the most part technically oriented and lacking broader management experience and training, tend to emphasized technical security measures that protect information against improper disclosure. Indeed, the term "security" is often (and incorrectly) assumed to refer exclusively to logical and physical access controls. Other risks, particularly service interruptions and fraud, may get little attention.

4) Senior managers do not consider the effectiveness of data processing security programs when evaluating the performance of middle managers, and responsibility for proper security usually is not explicitly included in management job descriptions. Without a credible QRA how can one evaluate the effectiveness of the security decisions taken by a subordinate?

5) Persons who are inexperienced in the conduct of risk assessments and unaware of QRA estimating techniques and sources of risk information, may believe that it is not possible to make credible estimates of risk factors, and conclude that QRA is not feasible.

The Importance of QRA to EDI Systems.

Since EDI systems cause Government agencies and private sector organizations to share the risks inherent in EDI systems, it becomes increasingly important for Government agencies to rationalize the risk management process. EDI systems can particularly benefit from the use of QRAs to manage risks because it is relatively easy to estimate the loss potential associated with financially oriented EDI systems. These are the major risk categories:

1) Service Interruptions. If an EDI system fails to process EDI messages timely, associated operations are delayed resulting in extra expenses, reduced productivity, lateness penalties, reduced or lost revenue, and delayed collection of funds. In most cases estimating the characteristics of interruption threats, and the cost impact of interruptions will be quite straightforward because of the "business" orientation of EDI systems.

Sabotage is a significant EDI risk. When an organization concentrates its essential business records in EDI systems and eliminates paper records, it faces an increased exposure to "software" sabotage by a disgruntled insider, or by an outside who seeks to gain an advantage over the victim. The possibility of massive data destruction and operational disruption exists. While it may be difficult to estimate the occurrence rate of sabotage, it should be relatively easy to estimate the cost impact, and generate reliable SOL estimates.

2) Fraud. Because many EDI transactions are performed automatically, computer-to-computer, without human oversight, the possibility of massive fraud exists. Experience has shown that some people are dishonest and will take advantage of opportunities to steal. Once a way to found to manipulate a EDI system, there may be little limit on the size of the manipulation because of the lack of human oversight. Consequently, there is a strong incentive for dishonest people to find ways to defeat EDI internal controls. Because loss potential can be related to the magnitude of the funds transferred between organizations, reasonable estimates of the cost impact and SOL can be generated even if fraud threat occurrence rates are difficult to estimate.

3) Information Disclosure. EDI systems commonly process personal and private information, as well as extremely important commercial data such as bid prices and proprietary information and processes. The greater the value of the stored information to an intruder, the greater will be the risk of unauthorized disclosure. Large scale EDI systems are likely to have relatively high disclosure risks.

Critical Aspects of EDI Risks.

A preliminary analysis of typical EDI systems suggests the following areas would particularly benefit from the use of QRAs:

1) Message Authentication. The cost of defeating an EDI system's authentication depending on the characteristics of the authentication system used. The

benefit depends on the functions performed by the system. The optimum authentication system will not always be obvious without the help of a QRA.

2) Use of artificial intelligence to validate messages, actions, etc., for reasonableness with a maximum probability of detecting fraudulent messages and a minimum probability of false alarms.

3) Use of artificial intelligence to detect attempts at unauthorized access to data, log appropriate information about the attempts, and alert system operators in real time.

4) Optimizing the frequency and scope of data file back-up. The optimum program is determined in part by the volume of data updates, the extent to which lost data can otherwise be recovered, and the expected frequency of recovery from back-up data. Because of the complexity of the relationships, the optimum program will not usually be obvious.

5) Optimizing plans for recovery from disasters, and business resumption planning. Determining factors are similarly complex in this area.

6) Optimizing financial, operational, and compliance auditing of EDI systems to deter and detect fraud. It is important to identify accurately the critical controls and procedures in order to select the best audit program.

Proposed EDI QRA Research.

To make the best use of QRA in supporting the design, implementation and operation of EDI systems, research is needed in the following areas:

1) Develop procedures for evaluating individual EDI systems to determine the appropriate QRA techniques to use, and develop these techniques. Since EDI systems vary widely in size, criticality, and vulnerability, no single QRA procedure can be appropriate for all EDI systems. Instead, several EDI QRA procedures should be developed, probably about seven to ten, that correspond to the models for a reasonable range of real-world EDI

systems from small and simple to large and complex.

2) To the extent feasible and useful develop standardized EDI parameters and estimating techniques for threats, loss potentials, and vulnerabilities that can be adapted to the individual QRA procedures.

3) Develop model policy and procedure statements, suitable for organizations of all sizes, both private and government, for the sound management of EDI risks. To the extent possible, define good security practice so as to simplify risk management.

4) Define and establish a suitable forum for the ongoing exchange of EDI risk management information.

5) Use QRA techniques to analyze the potential cost benefit of an independent service to log automatically messages between EDI partners as a way to resolve disputes. The analysis should include factors such as log record retention time, record detail, central vs. regional logging facilities, sample logging, off-line batch logging, and cost recovery.

6) Consider the relationship between the concepts of holder in due course, "prudent man" management, and other factors used to determine legal liability for damages caused by EDI system failures including fraud, and the extent to which the selection of security measures for a failed EDI system was supported by an adequate QRA.

Since both Government and the private sectors will benefit from this research, it is important to ensure that both communities are represented when specific research objectives are defined, and that research team members represent a broad range of EDI users.

Risk Management Objectives

1) Optimize the selection and implementation of security measures by minimizing the risk cost.

2) Protect against catastrophic risk losses. (Can't be cost optimized.)

Risk Management Definitions:

Risk

A sudden, damaging event, not normally planned for that has a material impact when it occurs. Often referred to as a threat.

Risk Management

Making rational provision for the occurrence of risks. Risk managers plan for risks. Line managers don't.

Risk Assessment

The process, formal or intuitive, whereby risks are identified and their expected losses are estimated.

Security Measure

Any device, procedures, or environmental factor that reduces the loss caused by a threat event.

Risk Cost

The sum of

the cost to install and maintain security measures

and

the losses caused by risk events.

The Effect of Security Measures.

The purpose of security measures is to reduce FUTURE losses.

There are four reasons for selecting a given security measure:

Automatic Selection.

1) It is required by law. (Exit signs.)

2) The cost is trivial but the benefit (loss reduction) is material.

Select IF Benefical. (Optional).

3) Its total cost is significantly LOWER than the REDUCTION in future losses it is expected to achieve. Reduces risk cost.

4) It mitigates the impact of a catastrophic threat.

The dilemma is that managers must make security resource allocations NOW to avert FUTURE losses.

What Determines Future Losses?

Future losses of a "system" are determined by these three factors:

1) The rate at which threat events occur.

2) The potential for loss inherent in the functions performed by the system and the assets of the system. (The worst case loss.)

3) The vulnerability of each function and asset with respect to each threat.

   Note: Vulnerability to a threat is NOT significant unless the threat is expected to occur. This is why a vulnerability assessment is NOT the same as a risk assessment.

The Need for Quantitative Estimates of Expected Loss.

The cost of security measures is measured quantitatively.

If the cost of a security is material, its benefit (reduction in future losses) must also be measured (estimated now) quantitatively in order to optimize its selection.

There are two kinds of loss estimates:

1) Annualized Loss. (Sometimes referred to as Annualized Loss Expectancy or ALE.) ALE is the expected loss expressed at an annual rate.

   ALE helps with selection of beneficial security measures, prioritizes functions for disaster recovery planning, and identifies major risks.

2) Single Occurrence Loss. The loss resulting from a single occurrence of a threat.

   SOL identifies potentially fatal risks.

Obstacles To Quantitative Risk Analysis.

1) Unwillingness to invest money to avert future losses. Reactive rather than proactive security planning. Spend limited funds on obvious (or popular) security measures.

2) Government managers do not routinely emphasize Return on Investment when planning expenditures. (Allocate budgeted amounts.)

3) MIS managers lack training and management experience with protection of assets as a basic management role.

4) Performance evaluations of managers do not consider effectiveness of security management.

5) Managers inexperienced with Quantitative Risk Analysis may conclude that credible loss estimates are impossible.

The Importance of QRA to EDI.

EDI risks are not obvious by inspection. EDI system managers (who make the security decisions) are not familiar with the impact of security failures on end users.

1) Service Interruptions. Losses depend on the character of the functions being supported by the EDI service.

2) Sabotage. Total dependence on electronic documents and record keeping creates major a vulnerability to insider "software" sabotage.

3) Information Disclosure. Electronic documents (including non-text documents) are relatively easy for insiders and outsiders to steal compared with traditional documents.

4) Fraud. Reduction in (or elimination of) human oversight increases the risk of delayed discovery of systematic fraud.

The benefits of EDI translate into new or enlarged risks.

Using QRA to Select EDI Security Measures.

Here are some examples of areas where QRA can support design and operating decisions:

Message authentication.

Message "reasonableness" validation.

Automated detection of "hacker" attacks.

Automated detection of fraud attacks.

Optimized data back-up programs.

Optimized disaster recovery plans.

Optimized compliance auditing programs.

Proposed EDI QRA Research Topics

1) Establish a set of (about five to seven?) EDI system models, ranging from simple to large and complex, and a corresponding suite of QRA techniques.

2) Identify and determine standard EDI QRA parameters.

3) Model policy and procedures statements, and, to the extent possible, define good EDI security practice.

4) Establish a permanent forum to exchange EDI QRA know-how.

5) Estimate cost/benefit of independent message logging service.

6) Relationship of security based on careful QRA and the concepts of holder in due course, and prudent-man management.

Presented by: Jim Orr - Director, Support Services
Blue Cross of California

A. Introduction

Thank you for the opportunity to participate in the Workshop on Security Procedures for the Interchange of Electronic Documents, sponsored by the National Institute of Standards and Technology (NIST). I appreciate the chance to express my views as to effective means for the secure exchange of information.

In this paper, I will begin by discussing what we have in common with many industries striving to increase their use of electronic interchange, rather than relying on paper-based communications. You'll find we have many common goals.

I'll then highlight what is unique about the health care financing business, and the security problems that uniqueness brings. I hope to demonstrate those problems so that our efforts in the workshop can include all situations, thereby making any solutions and recommendations we develop appropriate to all.

B. In Common with Other Industries

Like most industries, the health-care industry is using electronic transactions to a greater degree each day. The transactions being used are a combination of proprietary and the standard transaction sets being developed and approved by the American National Standards Institute (ANSI) Accredited Standards Committee (ASC) X12N subcommittee. In fact, there are more than 400 transaction formats currently in use. I'll say more about standards and the X12N committee later.

How important are electronic transactions to the health care industry? In the recent report from the Workgroup for Electronic Data Interchange (WEDI) to

Secretary of U. S. Department of Health and Human Services[1], it was estimated that among $4 billion and $10 billion could be saved annually with the use of Electronic Data Interchange (EDI). These administrative costs have a direct impact on the costs to the consumer (and to his or her employer) of health care insurance.

The savings are anticipated to be realized in the same way as other EDI implementations:

Increased accuracy
Reduced 'human' intervention
Improved claims turnaround
Better customer service
Etc.

Clearly, there are compelling financial incentives for health-care providers and insurers to increase the use of EDI. I would suspect a similar set of incentives exists in your area, as well.

One of the questions to be understood in our workshop is:

How does the implementation of EDI increase the security risks to either the sender or the receiver?

Some of the possible message communications risks are:

Injecting fraudulent messages
Wire tapping
Masquerading
Altering message data
Communications interruptions

Many of our security risks are similar to other industries. We have financial transactions, and, therefore, direct financial risk. In fact, every claim transaction may be thought of as a check request. Of course, we're different from other financial institutions in that, unlike a bank, a patient may

[1]Report from Workgroup for Electronic Data Interchange To Secretary of U. S. Department of Health and Human Services
July, 1992

'overdraw' his account, based on medical services rendered, without regard to the amount of premium dollars previously 'deposited'. This increases our risk to those who would submit fraudulent claims.

Like other industries, we have high volume of transactions. During 1990, the Blue Cross Blue Shield Association indicates that the participating plans received in excess of 940 million claims. For private business, about 60% of the claims from hospitals were electronic, while only 20% of the claims from doctors were electronic. For Medicare Part A claims, more than 75% were electronic.

We are geographically diverse. While here in Maryland, I could go to the nearest bank and reasonably expect that my ATM card would work, and I could withdraw some cash from my account in California. If I tripped and broke my leg and went to the hospital here:

Could they admit me, knowing (electronically) that I had coverage?
Could they submit the bills electronically?
Could they receive payment electronically?

Here's another example: An IBM employee goes to the doctor in San Francisco. The doctor submits the claim to Blue Cross of California, who pays the claim. There is then a settlement transaction with the Empire Blue Cross Blue Shield, because IBM's headquarters are in New York.

Here's the point - health care processing is not limited geographically, and the paths for electronic documents are *not* all predefined nor the subject of specific point-to-point trading partner agreements.

C. What's Different about Health Care?

As mentioned above, it is clear that we face the same risks as other industries as we use electronic transactions. For this workshop, however, please consider what is unique about health care and health care financing:

We are dealing with data concerning the medical status of individuals, with those individuals specifically identified.

So that we can properly pay the claim, we receive information about who's gone to a fertility clinic, who's been tested for HIV, who's gone to a drug rehabilitation facility, etc. I'm sure all of us understand the sensitivity of the data in these examples.

Here are some of the ways we treat health care information, thereby possibly increasing the risks of unauthorized access:

Self-Insured Accounts

Several of the customers of Blue Cross of California are self-insured. To that end, Blue Cross provides enrollment and claims processing services only, not underwriting. One of those provided services is a reporting of actual claims expense back to the customer; that is, we tell our customer, the employer, the details about the claims from their employees we have processed on their behalf.

We are careful to send only that information to the self-insured company that is necessary for their cash management and other financial needs. No procedure codes nor names of providers are sent.

This arrangement is unique as compared to the normal health care insurance arrangements in that we are now telling the employer about claims for sensitive medical treatments and tests. The fear is that the wrong use of this information could result in employment, salary and promotion decisions being made in light of the medical information. The security issue for us is:

How can we protect the privacy of the individual, while still providing the required financial information?

I believe that the same type of dilemma exists in the public - private exchanges of data.

Electronic Payments

As we proceed with the implementation of electronic payments to providers (hospitals, doctors), the issue of how to treat the Remittance Advice is being raised. That is, should we be sending through the banking system all the text necessary to describe how Blue Cross has processed one or more claims, resulting in a payment? The Remittance Advice can contain patient-specific information.

How can the receiving system merge and understand both the Payment and the Remittance information, when they are received at different times?

Too Much Data

For medical review purposes, we will sometimes ask a provider to send us more information about the medical situation concerning the specific claim - an operations report, lab test results, additional symptoms, if any, etc.

The provider may, then, anticipating the possibility of even more information requests, reproduce the entire medical record and send it to Blue Cross of California.

Of course, the first thing we do is microfilm it. We now have two copies (paper and microfilm) of data that is in excess of our needs with respect to the claim in question.

Two lessons here:

We need to provide information to the providers as to what we expect under what conditions.

Electronic patient records may make the exchange easier, including the disposal of unwanted information after receipt.

D. Our Interaction Today with Uncle Sam

Today, Blue Cross of California interacts with the United States government in health care as a Medicare processor. We have also bid on the contract to become the CHAMPUS processor in California and Hawaii. (As you may know, CHAMPUS is the program to provide medical coverage to dependents of military personnel.) These interactions are different from traditional claims processing that starts with the provider and ends with the payer. With Medicare and CHAMPUS, the electronic path of a claim continues on to the HCFA or DoD agency. I thought it would be interesting to look at some of the information requirements of those two agencies as an example of the public-private data exchange partnership now in place.

Medicare

If a provider were to query the Common Working File (CWF) of Medicare at one of the nine regional hosts, the response includes:

Date of Birth
Psychiatric Benefits (Current and Prior Year)
Hospice Benefits

I believe that an argument could be made that not all providers have a medical or financial requirement to understand that their patient has or hasn't received psychiatric care this (and last) year under Medicare. While it is important for the physician to treat the whole patient, and to understand all elements that may impact diagnosis, I believe that the purpose of this query is to display benefits information, not patient medical records. To that end, should not the response be tailored to the type of provider making the query?

Champus

Let's look at the data exchanged between the CHAMPUS contractor and CHAMPUS when an eligible dependent submits a claim:

First, there is an eligibility inquiry that uses the following data:

Sponsor's Name, SSN, Pay Grade
Patient's Name, SSN, Date of Birth
Not-Available Status Code

Example: A code value of 12 is for diseases of the male reproductive system

On the claim itself, we have:

Sponsor's Name, SSN, Pay Grade
Patient's Name, SSN, Date of Birth
Provider Taxpayer Number
Primary and Secondary Treatment Diagnoses
Principal and Secondary Operation / Non-Surgical Procedure Codes

Note that on an adjustment, CHAMPUS requires: "All data reported on the original [claim] must be resubmitted except for signed numeric fields, and those non-signed numeric fields requiring correction."[2]

These data exchanges are very important to ensure that CHAMPUS and the contractor are in synch concerning eligibility of the dependents. The exchanges are designed to help each agency understand the status of the sponsor (active, retired, etc.) and the dependent (student, name change, etc.) so that benefits can be paid accurately and in a timely fashion.

However, the requirements of CHAMPUS also causes the exchange of a large amount of possibly confidential data that may not be necessary each time. Without secure transmissions (or batch file transfers), the risk of unauthorized access increases with each exchange of data.

Given that exchange of data, what are the security requirements for the EDI transactions, along the entire path, including the data from the provider (doctor or hospital) to the payer (Blue Cross of California)? The above CHAMPUS example is why we are here at NIST.

E. Recommendations

For the purposes of our workshop, I have put together the following pages as examples of how we might develop categories and levels.

[2] CHAMPUS Automated Data Processing and Reporting Manual, Office of Civilian Health and Medical Program of the Uniformed Services, Aurora, Colorado, 1-90. Page 1-12.

For levels of risk, I have selected only three: High, Medium, and Low. Intuitively, three seems to be the maximum number of ways I thought I could slice the risk.

For categories, I have used the four specified in the announcement of our workshop: Authenticity, Confidentiality, Integrity, and Timeliness. There is plenty of blank space for you to fill in examples of your own, or to add comments or questions.

Category: Authenticity

Examples of Low Risk Level:

- Basic Eligibility Inquiry
- New Member Enrollment
- Any Claim less than $N

Examples of Medium Risk Level:

- Claim Status Inquiry

Examples of High Risk Level:

- Pre-authorization Response
- Any Claim $N or more

Category: Confidentiality

Examples of Low Risk Level:

- Any transaction where the data does not identify an individual

Examples of Medium Risk Level:

- Transactions involving rates (membership or provider)

Examples of High Risk Level:

- Any transaction where the data does identify an individual

Additional Comments:

Confidentiality may only have two states - the data is either private or not (unlike the service, there may be no Private and Private First Class).
Note that the laws for confidentiality differ from state to state. You may want to refer to the WEDI report[3] for a detailed discussion of this issue.

[3]Report from Workgroup for Electronic Data Interchange To Secretary of U. S. Department of Health and Human Services
July, 1992

Category: Integrity

Examples of Low Risk Level:

- Eligibility Inquiry
- Claim Status Inquiry

Examples of Medium Risk Level:

- Batch Claim File Transfer
- Imaged Documents such as X-Rays

Examples of High Risk Level:

- Electronic Claims or Premium Payments
- Electronic Patient Record

Category: Timeliness

Examples of Low Risk Level:

Claim Status Inquiry

Examples of Medium Risk Level:

- Electronic Claim or Premium Payment (depending on delay and $ amount)

Examples of High Risk Level:

- Eligibility Inquiries for Admittance
- Pre-Authorization Responses

Given the above categories and levels, how does one determine where a specific electronic document fits? That is, how does one 'map' a document onto these categories and levels?

One technique used by the ASC X12N (Insurance) Security Work Group is that of providing a self-scoring security evaluation questionnaire. I have included a portion of the current version in Appendix A. This technique, along with the handbook that indicates what security techniques to use for which type of risk, allows for the evolution and implementation of new needs and new techniques. Please review Appendix A to get an appreciation of this approach.

F. Other Recommendations

What else can we do during this two-day workshop?

As a group, we can:

1. Develop a recommendation for all government agencies that less data in general and specifically less individual-identifiable data be required from the private sector.

As indicated above, there is a large amount of medical data that can be sent as the result of a patient going to a doctor. Are all data being required absolutely necessary? Are there other techniques by which the data about the medical encounter and the data about the patient can be sent separately? Would actuarial requirements be met if the patient were identified generically (age, sex, etc.) without the name and SSN?

2. Develop an action plan to define: "In EDI, who is responsible for risk analysis?"

What is 'reasonable' depends on the risk. In a traditional trading partner arrangement, is it the sender or the receiver who evaluates the risks associated with the electronic transactions? What about in an 'open' system where the path of the transaction and the reply may not be predetermined, and may pass through third party services providers?

I would challenge the statement on page 5 of our workshop announcement that states:

> "It will remain the responsibility of each Federal agency to specify the particular level and category of each electronic document that it interchanges."[4]

That statement does not establish consistency among various government agencies for the security of the same data, nor does it deal with the various state laws in which the private sector must deal.

It also could lead to inconsistent requirements among agencies. For example, HCFA could specify one category and level of security for electronic claims, and DoD, through CHAMPUS, could specify another category and level. That could mean that for the same provider for the same type of medical treatment, there could be two security requirements for the transmissions of those claims: one set of requirements from the provider to the payer, and another set from the payer on to Medicare or CHAMPUS.

---

[4]Announcement of a Workshop on SECURITY PROCEDURES FOR THE INTERCHANGE OF ELECTRONIC DOCUMENTS, Computer Systems Laboratory, National Institute of Standards and Technology, United States Department of Commerce. March 24, 1992.

3. Develop acceptance criteria for the certification of encryption techniques and tools.

Rather than having NIST or other agencies select one preferred algorithm, why not develop the requirements or the criteria by which encryption products could be certified? In this manner, a number of products could be deemed acceptable, thereby enhancing the implementation choices.

In reviewing the UN/EDIFACT documentation on security[5], they make provisions for the following encryption techniques:

Message Authentication Code
FEAL-MAC
DES Modes of Operation (Cipher Block Chaining and Cipher Feedback)
Hash functions using a n-bit cipher algorithm providing a single-length hash code
Hash functions using a n-bit cipher algorithm providing a double-length hash code
Square-mod n hash function for RSA
Modification Detection Code -- (IBM System Journal)
BGC-7.1 hash function
MD4 Message Digest algorithm
MD5 Message Digest algorithm
Secure Hashing Algorithm
Mutually-agreed.

The point here is that certification could help keep the regulations concerning encryption current with generally-accepted technology.

4. Develop a recommendation that each agency implement and publish specific guidelines with respect to definitions of sensitive data.

As mentioned above, one of the WEDI recommendations for dealing with confidentiality is Federal, pre-emptive legislation to eliminate the differences in regulations among the individual states. In that same manner, each Federal

[5]Recommendations for UN/EDIFACT Message Level Security, from the UN/EDIFACT Security JWG, May, 1992 Version

agency could implement specific guidelines that define the security requirements of each transaction. These guidelines would provide the 'rules' for EDI, against which all participants could be measured.

5. Develop a recommendation that signature requirements for documents be minimized, so that electronic signatures can be used in lieu of paper documents that require signatures.

This approach would make the use of EDI more attractive, and less limited by the use of paper that may be required today to accompany the electronic transmission.

6. Develop a recommendation for an approach for a National Health care ID card.

The technology has been proven with ATM applications. A national ID card could include PIN-code techniques, thereby increasing the security by helping to verify the holder of the card.

One approach would be to participate in the Patient ID Card Group of the X12N Insurance subcommittee.

Are there other industries to which an ID card would apply?

7. Develop an action plan that would utilize the current standards-setting committees within X12 or other organizations.

On the basis of the kind of standards required, NIST can participate in the appropriate committee to continue to have influence and to receive a broad base of input from representatives of many organizations.

On the Optimal Expenditure of Computer Security Costs

by

Roy G. Saltman
National Institute of Standards and Technology

Note: The views expressed in this paper are the author's, and not those of the National Institute of Standards and Technology.

The question of how much security is enough is not a question that began with computers. The issue is pertinent in any situation of defense against an unpredictable opponent, regardless of whether the adversary has human intelligence and effective technology, or is a natural force such as a hurricane or earthquake.

In many situations, it is not easy to quantitatively specify levels of expenditures, or to authoritatively articulate the solidity of defenses to be implemented, because the probabilities and strengths of attacks cannot be accurately predicted nor the losses accurately estimated. In some of these cases, guidelines and general principles are offered instead of specificity.

General Principle for Federal Agencies

A general principle is enunciated in the Computer Security Act of 1987 (P.L. 100-235). Under this act, each Federal agency is to establish a plan for computer system security and privacy that is

"commensurate with the risk and magnitude of the harm"

resulting from any compromise of system integrity or confidentiality. Each plan is to be submitted to NIST for its advice and comment. The plan is expected to involve expenditure of resources, since it "shall be subject to disapproval by the Director of the Office of Management and Budget." One may interpret the principle as stating that there must be a trade-off between expenditures for security and the cost that a loss would cause, given some understanding of the likelihood of the loss.

Principle Available to the Private Sector

In the private sector, an authoritative source discussing the implementation of security is the legal document called Title 4A of the Uniform Commercial Code (UCC). The UCC was developed as a proposed law to be adopted uniformly by the States, to simplify interstate commerce. If State commercial codes differed, interstate commerce could be a morass of conflicting legal requirements. Title 4A, concerning funds transfers, was drafted by the American Law Institute and the National Conference of Commissioners of State Laws. It is now adopted in most States, e.g., Maryland [1].

Title 4A, sec. 202, discusses liability for loss when a "security procedure" that is "commercially reasonable" is used by a customer to issue a payment order to a bank. "Commercial reasonableness" is not defined technically in Title 4A, but it is stated to be a question of law determined by, among other things, "the size, type,and frequency of payment orders normally issued by the customer to the bank, ... and security procedures in general use by customers and receiving banks, similarly situated."

However, what is in general use may not fully characterize what is prudent. Michael Baum, one of the leaders in the interpretation of laws applicable to EDI, cites the case of the ship T.J. Hooper. The ship foundered in a storm at a time when weather radios were available but not mandated to be used. The ship did not have a radio to receive weather alerts. Owners of lost cargo won a judgment against the ship owners, despite the lack of the requirement. As Baum states,

> "Regardless of current industry practices, if new or improved technologies and procedures are available at a cost lower than the value of the potential risks, failure to implement them could result in increased liability exposure." [2]

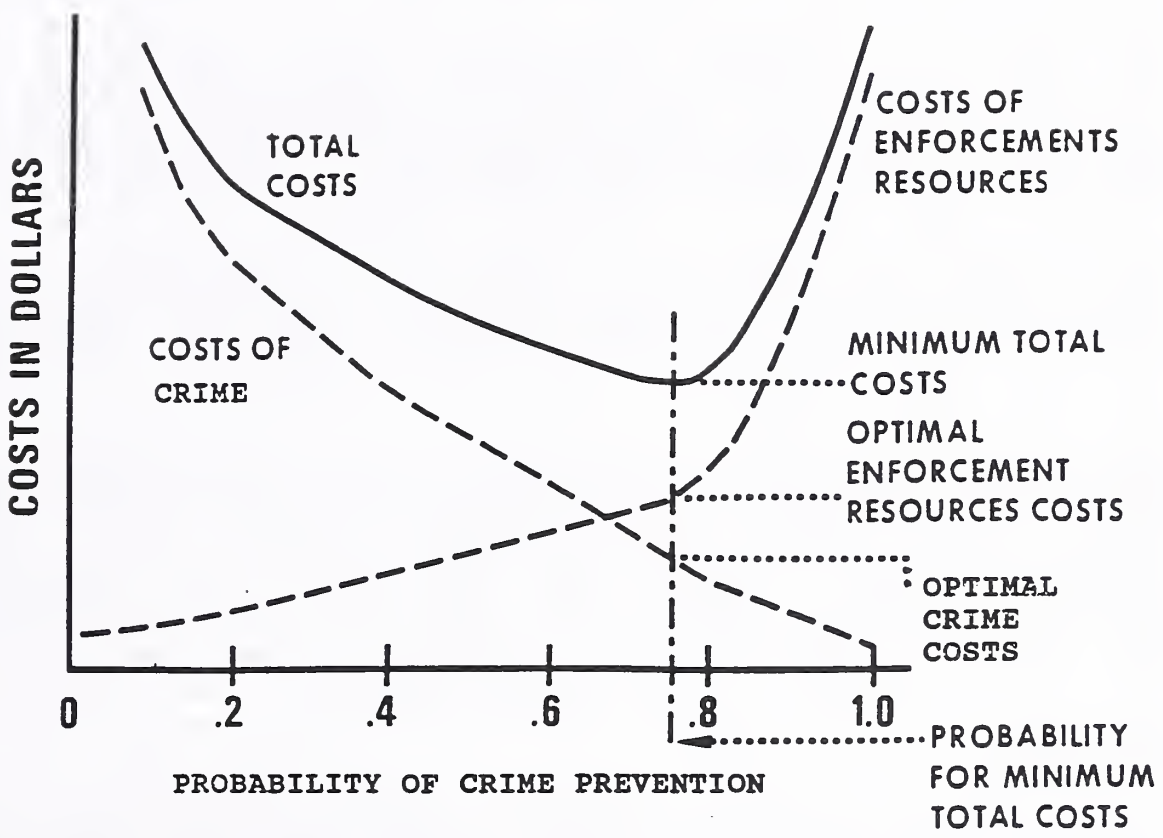
A key phrase, for the purpose of this paper, is "available at a cost lower than the value of the potential risks." Here, as in the Computer Security Act, a trade-off for security expenditures is proposed. Baum makes this concept explicit in a comment on those security experts advocating procedures and technologies that may result in extreme system trustworthiness but that may be too stringent for commercial use. Baum states:

> "While this position [of advocating stringent techniques] may be commendable in its intent, it does not adequately address realistic trade-offs between EDI security considerations and economic constraints." [3]

A Mathematical Interpretation

Economist Edwin Mansfield rhetorically asks "What proportion of the people who commit crimes of a certain kind should society try to apprehend and convict?" He continues, "Your first reaction may be that they all should be caught and convicted, but if so, it is easy to show that you ought to reconsider." [4]

Mansfield has presented a useful graphical explanation of the fact that it does not pay to enforce the laws perfectly. The explanation is contained in the figure included with this paper; it could apply to any type of crime, but it is used here for violations of computer security.



The Optimal Level of Computer Security Enforcement

It can be seen in the figure that total costs are the sum of two primary costs. One primary cost is the cost of enforcement of the laws. As more resources are utilized for enforcement, the probability of crime prevention increases. This result is due to one or both of the causes that (1) commission of the crime is more difficult technologically, or (2) apprehension is more likely based on the experience of other perpetrators.

This effect of increase of prevention assumes that there is a monotonic relationship between expenditure of resources and effectiveness of efforts. This may or may not be true initially, but if it is not, new procedures and technologies that are more effective at lower cost are likely to be found and used. Thus, it is likely to be true in the long run.

The second primary cost is the cost of crime itself. As the laws are enforced more leniently, and the probability of prevention decreases, the cost of crime increases. An assumption here is that the cost of crime monotonically increases with its incidence, an assumption that would be difficult to dispute.

Thus, total costs are the sum of two other costs, one of which correlates with an increase in prevention probability, and one of which correlates with a decrease in the same variable. Consequently, total costs have a minimum at some probability of prevention less than one. In effect, it does not pay to stop all crime.

This graphical presentation that there is an optimal expenditure of enforcement resources is consistent with the several similar concepts identified above. These concepts are that security implementations should be "commensurate with the risk and magnitude of the harm," (Computer Security Act), that security should be "commercially reasonable," (UCC 4A), and that procedures and technologies should be used that are "available at a cost lower than the value of the potential risks" (Michael Baum).

References

[1] The Annotated Code of the Public General Laws of Maryland - Commercial Law, Enacted by Chapter 49, Acts 1975; 1992 Replacement Volume. Charlottesville, VA, The Michie Company.

[2] Michael S. Baum, "Commercially Reasonable Security: A Key to EDI Enforceability," Actionline, Vol. 8, No. 11, Nov. 1989; Southfield, MI, Automotive Industry Action Group.

[3] ibid.

[4] Edwin Mansfield, Microeconomics, (Shorter Second Edition), W.W. Norton, 1976, pp. 146-147.

The Legal Viability of Electronically Submitted Environmental Compliance Reports*

David S. Schwarz, Chief
Information Policy Branch
Office of Policy, Planning and Evaluation
United State Environmental Protection Agency

*The views expressed in this paper are the author's and do not reflect any official position of the USEPA

I. INTRODUCTION

The USEPA is currently pilot testing the use of electronic data interchange (EDI) for industry submission of compliance reports under regulatory programs that implement air, water, and hazardous waste statutes. To date, for purposes of legal documentation, these pilots have all relied on concomitant paper submissions, e.g., a paper certification form. Hence, as yet the EPA experience with EDI does not offer a case that would exemplify true, paperless, electronic reporting, or would provide a test of its legal viability. The focus of this paper will therefore be prospective: considering the conditions that these pilot applications **will** have to satisfy before true EDI can be legally viable.

Our pilots are introducing EDI into regulatory programs that represent a substantial portion of EPA's core command and control responsibilities, and under these programs compliance reporting is used (in some cases together with inspections) as the primary mechanism for enforcing environmental standards and regulations. Of course, where violations are alleged, these enforcement activities culminate in some sort of legal proceeding--often in court--to seek civil and/or criminal penalties. It is in such proceedings that compliance reports play their most critical role: either to serve directly as evidence of environmental wrongdoing, or else--where there are issues of falsification--to demonstrate fraud in conjunction with other, independent evidence of environmental violations. Clearly, electronically submitted versions of compliance reports must be admissible as evidence in such cases with exactly the same force as their paper counterparts; for all practical purposes, this requirement provides both **the necessary and the sufficient conditions for the legal viability of EDI** in our applications.

Admissibility of a submitted document as legal evidence can be analyzed as entailing one or more of a number of standard elements: **originator authentication**, **non-repudiation**, **message integrity**, and **date/time certification**; it may also involve questions of **confidentiality**. For each of these elements in turn, the discussion that follows will attempt to identify the hard cases (if there are any)--that is, the kind(s) of environmental enforcement case(s) in which the element is most difficult to satisfy, and what such satisfaction requires. When we put these pieces together,

they should give us a picture of the strongest conditions that an environmental EDI application would have to satisfy. Against this background, we can then consider how strong the conditions must be for legally viable EDI in the typical or 'baseline' case of environmental compliance reporting.

II. REQUIREMENTS FOR ADMITTING DOCUMENTS IN ENVIRONMENTAL ENFORCEMENT PROCEEDINGS: THE HARD CASES

1. Originator Authentication. A legally valid compliance report must identify both the submitting organization--usually a private-sector company--and a responsible individual within that organization who certifies that the data submitted on the report is correct. In the normal case, these identifications are not an issue for enforcement officials. Either the data in the report show a violation or it does not. If they do, then appropriate follow-up actions are taken, e.g. fines are imposed. If there is a dispute at all, then it is over what the data show, or whether they are accurate, and this is resolved by carrying out plant inspections, looking at plant records or other data, or whatever.

However, establishing the identity of a certifying individual can be a concern where the Agency is trying to assign liability for purposes of obtaining damages. These are typically cases of civil liability, and the concern will be particularly pressing where the company in question is insolvent--so that any liability that exists passes to responsible individuals. Probably the best examples come from the Superfund program, in determining "potentially responsible parties" (PRPs) for a Superfund site cleanup. Often, the company that created the problem has gone out of business. The task is then to trace individuals from that organization who were responsible for the activities--e.g., the disposal of wastes--that now necessitate a cleanup. To identify such individuals, it is often very helpful to be able establish that it was their signature on a hazardous waste manifest, or that they were the certifying official in some financial transaction. However, this kind of evidence--connecting individuals with documents--may not be critical. These individuals may be assigned liability simply by virtue of the positions they held in the defunct company, and we can usually find evidence of this that is independent of their connection with the documents in question.

The identity of the submitting organization can also be an issue, for example, where there is some question that this identity has been misrepresented. This would be a case of **fraudulent** submission, and, this fraud itself might becomes the subject of an enforcement action. As an example, consider a case where a company is supposed to have effluent samples analyzed by a certified laboratory, which is to submit a report of the analytic data directly to EPA (or to a delegated state agency). Fearing that an analysis will reveal serious violations of permit limits, the company might file such a report itself, fraudulently representing the report as coming from its laboratory. In establishing this fraud, the Agency must be able to demonstrate that the submitter is

in fact the company, notwithstanding the laboratory name and address and forged signatures on the report. Of course, in doing this, the Agency would not be relying on any connection between the submitted document and the names and signatures on that document--since it is just this connection that is being discredited. Indeed, the Agency might not need to rely on any evidence internal to the document or its circumstances of submission at all (although, of course, such evidence might be very useful), since there will also be the evidence of testimony, company records, and authentic laboratory analyses.

In the examples so far, either it's not critical that we establish the identity of a document's originator, or it's not critical to use the evidence provided by the submitted document itself to do this. Nonetheless, in at least one class of cases originator authentication will be of paramount importance, where we are taking **criminal** enforcement action against an individual. Probably the best example here is the case where a company submits fraudulent data under its own name, with the Agency then taking criminal action against the certifying officials, to penalize these individuals as responsible for perpetrating the fraud. To do this, enforcement officials must determine that the identity of these **individuals** is as the submission represents them, in addition to confirming that the submission did indeed come from the company in question, and was not just a malicious prank. After all, it is just in their capacity as certifying officials that these individuals are criminally liable. If it turns out that it is not their signatures on the document, then the case against the individuals in question is considerably weakened--unless it can be shown that they directed others to sign the document as their agents. It's worth adding that allegations of fraud may not be necessary to have a criminal case against individuals in their capacity of certifying officials. A submission may truly report certain environmental violations and might thereby make the certifying officials criminally liable, by virtue of providing evidence the these officials knew of--but did nothing to prevent or mitigate--the occurrence.

In any event, it is in these criminal cases that admitting a submission as evidence is likely to depend most heavily on determining the identity of its originator. And, in making this determination in such cases, some sort of demonstrable causal link between originator and submission is likely to play its most critical role. This will be clearer in the context of **non-repudiation.**

2. Non-repudiation. The issue of repudiation will be most important--and likely most difficult--in those cases where we are most concerned with originator authentication. As we have seen, these are cases of criminal enforcement against individuals, the most prominent being those that involve **fraud.** In these cases of criminal fraud, of course, the Agency must establish **intent** to certify falsely, and it is just this makes the question of repudiation must critical: the accused will obviously have strong

motivation to deny any intentional connection with the document in question, and the fact of a signature may be our only real evidence of anyone's intent.

In such cases we may, for example, have to refute claims that the signature was 'forged', or reproduced without authorization. To do this, normally only a demonstration of the necessary causal connection between certifying official and the document in question will suffice--e.g., demonstrating that given the style of writing and the chemistry of the ink, the signature could not have been produced by any other hand. It will not do to argue that by virtue of his/her position in the company, the official could be 'presumed' to have certified the submission, since this does not establish the requisite intent. Of course, there are sometimes other, independent ways of establishing intent, e.g. evidence of other actions of the individual that show knowledge of compliance problems and a disposition to deceive, or testimony of others that were authorized to sign for the individual under investigation. If sufficiently strong, these may allow prosecution to refute the attempted repudiations without appealing to the characteristics of the document itself. However, bound up as it is with establishing intent, it is hard to see that the issue of repudiation could be simply side-stepped--at least in these cases of criminal fraud.

Of course, as we turn from the criminal to the civil context, the question of intent drops out, and, correspondingly, any attempt to repudiate the identity of individuals certifying particular documents is much less of a worry. For example, while it may help greatly to know that John Smith's name was on a hazardous waste manifest, he may be liable for the unhappy fate of one of his company's barrels of PCBs simply by virtue of being the owner or operator of that company. Of course, we must still be able to refute attempts to repudiate the identification of John Smith's company as the originator, but this again may not require that we focus on who signed the document. Since a hazardous waste manifest involves many interactions between an originating company and other organizations, there may be many different ways to establish where it came from.

3. <u>Message Integrity.</u> For both originator authentication and non-repudiation, the question of fraud has divided the difficult cases from the relatively easy. We will find the same to be true here. Where fraud is not an issue, then generally an enforcement case will raise questions of message integrity where the defendant disputes the content of submissions that show violations. Such disputes do not normally create serious problems. Where the EPA receives reports that show violations, investigators typically follow up by conducting inspections and performing tests to provide confirming data before taking a case to court. It would be an extraordinary case indeed where the Agency's primary evidence of violations was signed certifications to that effect; hence, it is normally beside the point here to argue at length over whether what the Agency offers as the defendant's submission is in fact what the defendant submitted.

Of course, the Agency cannot be indifferent to the integrity of such submissions. A consistent pattern of discrepancies between the contents of Agency databases and what submitters claim to have sent would seriously undermine EPA's credibility in court--let alone its general ability to manage its programs. However, good audit trails and other QA/QC procedures associated with normal database management are usually regarded as adequate to take care of these concerns.

We are back to cases of fraud, then, and in this context the cases of concern are primarily those where we need to establish submission of false data. The issue of message integrity arises specifically where we must confront that claim that the appearance of falsification was introduced by errors in transmission or by changes made to the document while it was in the Agency's custody. It is not clear how critical this issue will be. On the one hand, if the allegation of falsification is confined to a single document, with no other supporting evidence of criminal intent--and against a background in which the defendant has generally been a good actor--then we can imagine that the government's case may turn on whether the document provided to the court can be shown to reflect exactly the data submitted, and the evidence here would have to be unimpeachable. On the other hand, it is difficult to imagine that EPA would bring criminal charges where the case hung by such a slender thread. Presumably, the Agency would have the evidence of actual environmental violations--which the submission in question was alleged to be trying to hide--as well as the confirmation offered by company records and by the testimony of witnesses. These could easily make a strong case for motivation to deceive, and against this background it would be difficult for a defendant to gain much credit for the assertion that the document in court does not truly reflect his submission. Of course, EPA would still need to demonstrate that it had adequate procedures in place to assure the integrity of its database--and the mechanisms for receiving input. After all, our claims of fraud in these cases is only as strong as our assertions of message integrity.

4. Date/time Certification. While EPA reporting requirements include specifications of schedules and deadlines, the failure to meet a deadline per se is rarely the subject of an enforcement action. However, a consistent pattern of late submissions, or a late submission connected with a critical violation of environmental standards will certainly be a part of the evidence that EPA would want to use on behalf of civil or criminal actions. Therefore, at least the date and time of the **receipt** of the submission would need to be verifiable in such cases. It is not clear whether we would need more than this, e.g. in cases where a dispute arose over when the submission was **sent**. If the role of date/time evidence was to show **patterns** of behavior, then a set of well-documented receipt dates should be sufficient to turn aside claims that there was an inherent inconsistency between when the Agency was claiming it got reports and when they were actually submitted.

It is worth noting, however, that there is at least one kind of case where the date of sending a document **is** of independent interest. Under hazardous waste rules, the liability of treatment, storage, or disposal facilities (TSDFs) for waste that they reship can depend on how long they have kept this waste; generally, if they keep it more than fifteen days then they acquire the generator's liability for the waste if they ship it elsewhere. The primary pieces of evidence for how long a shipment has stayed at a TSDF are the dates on the manifests indicating the arrival and the subsequent departure of this shipment--and these should coincide with the dates on which the TSDF sends copies of these manifests to the appropriate state agencies. Many legal consequences can flow from assigning generator liability for this shipment to the TSDF. And, it is at least possible that this attribution will turn on determining the dates on which manifests were submitted (although evidence from transporter records and testimony should also be available in the normal case).

5. Confidentiality. For purposes of admitting documents as evidence in enforcement actions, maintaining the confidentiality of what is submitted is not very relevant. Nonetheless, the ability to maintain confidentiality is of great importance to the Agency where submissions are legitimately asserted to contain 'confidential business information' (CBI). CBI constitutes a significant proportion of the submissions to the Agency under certain programs, especially those in the areas of pesticides and toxic substances. Under these programs, companies may submit closely held chemical formulae or formulator recipes, as well as sensitive production and financial data. Whether or not such submissions came to be associated with an enforcement action, failure to preserve the confidentiality of CBI--either as the data was submitted or as it was maintained--could subject the Agency to claims for damages. Therefore, quite independently of enforcement considerations, EPA will need special measures to assure the confidentiality of CBI data when submitted electronically.

III. APPLYING EDI AT EPA: THE HARD CASE AND THE BASELINE

Based on the preceding discussion, it appears that cases of criminal fraud erect the highest barriers to the admission of submitted documents as evidence. They require us to establish the identity of both the submitting company and the certifying individual(s). To refute attempts to repudiate these identifications, these criminal fraud cases generally require that we demonstrate a strong causal connection between the document and the certifying individuals--with evidence at least as strong as that which links a signature on paper to the hand that produced it. In addition, they require that we establish the integrity of our records or database--and our procedures for receiving submissions--with a high degree of confidence, and they may also require that we provide good evidence of the date/time of the submission.

Criminal fraud, then, provides the hard case for applying EDI

to environmental reporting. Compliance reports that are vulnerable to fraud will therefore place very strong conditions on an EDI application if the electronic submissions are to be legally viable. The question, then, is how far are such compliance reports from the baseline of typical EPA reporting. The answer, unfortunately, is that they are **at** this baseline. Looking across EPA programs, compliance reporting is almost wholly a matter of self-monitoring (or auditing) and certification. Clearly, a company finding itself in serious violation of environmental standards, or limits set by permit, has at least some potential motivation for misrepresenting its situation--in hopes that this will go unnoticed by environmental agencies and that the company will thereby escape the penalties. Across EPA programs, then, the threat that such fraudulent reports will be uncovered, and subject their authors to criminal penalties, is probably the keystone of our enforcement strategy.

The possibility of fraud is certainly of concern in the cases of the reports for which EPA is currently attempting to introduce EDI: the hazardous waste manifest (HWM) and the discharge monitoring report (DMR). Taking these in order, the HWM is used, among things, to identify the contents of (usually sealed) containers that comprise the shipment of waste. The generator of the waste prepares this document, and--if he is trying to get rid of materials that are very difficult and/or expensive to have disposed of legally--he may be tempted to misrepresent what he is shipping. EPA and its delegated State agencies would certainly want to be able to deter this sort of fraud with the threat of criminal prosecution.

Turning to the DMR, companies that have permits to discharge effluent under the National Pollutant Discharge Elimination System (NPDES) are required under the terms of their permits to sample and test their effluent at specified intervals to assure that they are conforming to the required pollutant limits. The DMRs report the results of such tests, and, obviously, if such results show serious violations of permitted limits then there is the temptation to try to disguise this fact. Since a DMR may give EPA the only warning it could have of an impending environmental catastrophe--e.g., the 'killing' of a body of water--we cannot treat lightly the possibility of fraud in this case either.

Between them, the DMR and the HWM very likely affect the overwhelming majority of the organizations regulated under environmental statutes--a universe that numbers well over 100,000 entities. Therefore, it does not seem possible to restrict the scope of the 'hard' cases for legally viable EDI by drawing a line around a **subset** that might be involved in criminal fraud. This is probably the place to scotch another hopeful thought as well: that we might draw a line based on the dollar value of damage or risk. Unfortunately, the dollar value of environmental damage that might be associated with fraudulent reporting cannot be predetermined. It is not like the case of a contract or purchase order where we know and specify in advance that we are dealing with a quantity of

goods or services worth less than $ 25 K. So far as we know, **any case of misreporting could potentially entail millions of dollars of damage; we have no way of excluding this possibility in advance.**

IV. CONCLUDING THOUGHTS

To summarize this analysis, in applying EDI to environmental compliance reporting we must generally provide for the contingency that any particular submission may have to satisfy conditions for admissibility in criminal fraud cases. This would seem to require, among other things, the general use of an 'electronic signature' that creates a link to individuals originating the submission strong enough to support a claim of criminal intent. It likely also requires strong measures to validate the integrity of our EDI transmissions, and, possibly the application of something like 'write once read many' (WORM) technology to assure the integrity of our databases. In addition, we may need an audit trail that certifies date/time of transmission, and provisions for confidentiality of message--such as encryption--if we are to receive CBI.

Many of these requirements would be taken care of by the infrastructure that EPA might in any case expect to have in place to implement EDI, e.g., secure databases and appropriate value-added network (VAN) services. However, it is at least possible that provisions for adequate 'electronic signatures' would also impose special technology requirements on our rather large community of submitters. In wrestling with this problem, it is likely that EPA will not be alone, since our concerns with criminal fraud in compliance reporting appear to have strong analogues in the context of filing tax returns. As the Internal Revenue Service (IRS) moves toward paperless tax filing, they, like EPA, will need to preserve the use of their submissions in criminal prosecution of fraudulent reporting. If there are workable approaches to this problem for the IRS, these may point the way toward (or provide) solutions for the EPA.

AUTHENTICITY AND ASSURANCE

Dr. Horton Sorkin, Howard University

1. Scope:

This paper addresses the choice of EDI authenticity scenarios including nonrepudiation with concerns regarding 1.risk, 2.costs, and 3.availability. This discussion will be limited to current technology. It also addresses available authentication and nonrepudiation mechanisms that can be used as a result of the currently available X12 security standards and access control technology that can exist with the network provider. Important security issues that are not covered include crypto key management, records management, and telecommunications management. Privacy and confidentiality, system availability, and message contents integrity are not directly discussed.

2. Laws, Regulations, and Liability:

The assumption is that the parties are working in a "green" light environment. Due professional care and documentation can result in expert testimony. Therefore problems can be resoved in regard to auditor trails, insurance claims, or litigation.

3. Controls, Vulnerabilities and Threats:

Controls are those system features that protect against potential threats that may result in losses. This approach to motivating the use of controls is part of the international standard:

"In a data communications environment, the term 'security' is used in the sense of minimizing the risk of exposure of assets and resources to various vulnerabilities. A vulnerability is any weakness of a system which could be exploited to breach the security policy of the system. A threat is a potential violation of the security of the system.[1]"

Controls can mitigate vulnerabilities. Controls must be designed as part of the system and be operating. The auditor can evaluate the potential threats to the security of a system by determining what controls are not present or the system's vulnerability.

THREATS to SECURITY

The first step to control design is to ascertain what could go wrong. What could go wrong (threats) is that desired security goals are not attained (at a reasonable cost). Errors and irregularities can compromise system security. The standards address security by emphasizing security threats because of irregularities. The major threats[2] to desired message authenticity attributes are:

identity interception: the observation for misuse of the identity of a user involved in a communication for misuse,

manipulation: the deletion, insertion, misordering, or replacement of data during a communication by an unauthorized

---

[1]. ISO/IS 7498/2 page 41.

[2]. ISO/DP 8594-8, Annex A.

person,

masquerade: the pretense to be a different user to gain access to information or acquire unauthorized system privileges,

replay: the recording and subsequent replay at some latter time of a communication,

repudiation: the denial of an actual user of having participated in a communication,

CONTROLS or SECURITY MECHANISMS

Controls (mechanisms) are available to contain the threats. The threats are contained because the system is made less vulnerable with the addition of controls. The controls that can provide for authenticity, integrity, and confidentiality are access control, authentication exchange, data integrity control, digital signature, encryption, notarization, routing control, and traffic padding.

SECURITY ATTRIBUTES PROVIDED BY THE MECHANISMS

This section discusses the security attributes that may be desired to mitigate vulnerabilities. Again we are taking as our primary attributes authentication (with access control and non-repudiation).

authentication: basically this attribute involves a form of acknowledgement or authorization that the transaction should take place. It takes two forms; Peer Entity Authentication and Data Origin Authentication. Peer entity provides verification that the party communicating with the receiver is the one that it claims to be. Data origin is a method of checking that the data transmitted is what it is supposed to be. Both of these checks are provided through the authentication exchange, which is the mechanism that checks the identity of the sender and corroborates the data. This attribute provides a control against the message's being sent incorrectly, message modification and an unauthorized person trying to gain access to the system.

access control: this attribute is designed to prevent the use of the system by unauthorized personnel or in an unauthorized manner. A company would not want its telecommunication system to be tapped into and used to order and reroute merchandise or other assets to some enterprising hacker.

nonrepudiation: the two nonrepudiation forms concern the originator and receiver. With originator non-repudiation the receiving party gets proof that the data is coming from its asserted origin. Should the sender try to deny dispatch after giving such non-repudiation, this provides the proof that the message had been acknowledged previously.

Receiver non-repudiation would involve acknowledgment by the destination entity that the message was received. Likewise should this party later try to deny that the message had been received, it will be difficult to deny.

Non-repudiation is a valuable attribute because it provides controls against a set of threats: incorrect dispatch, lost messages, incorrect messages, and modification of messages.

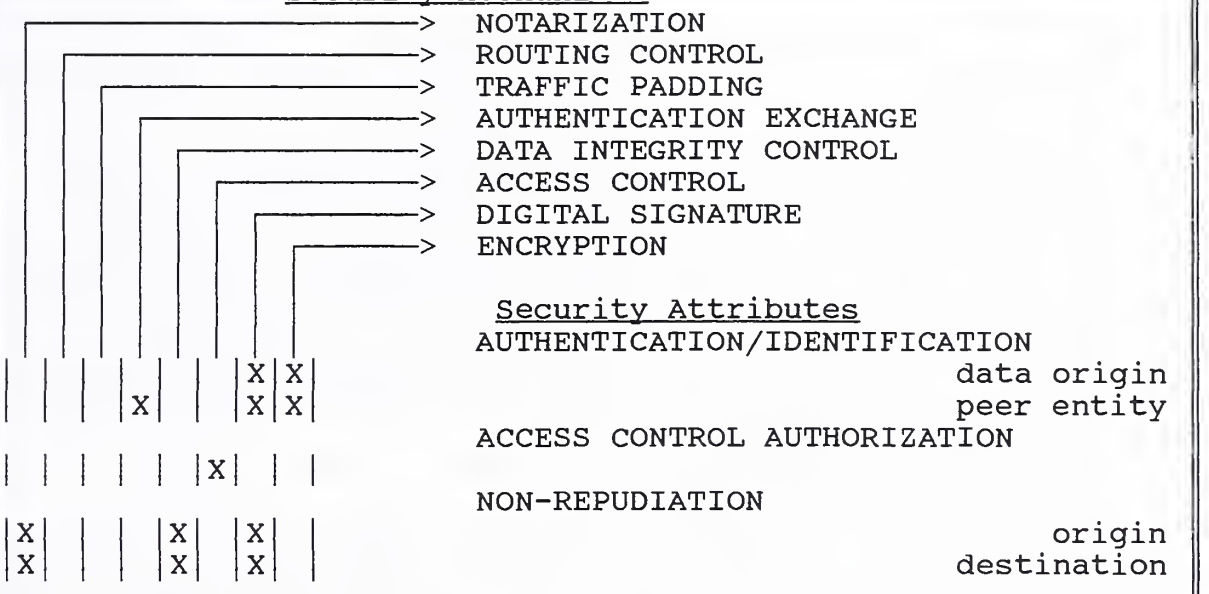
RELATIONSHIP of SECURITY CONTROLS and ATTRIBUTES

The choice of appropriate security mechanisms results from the attributes desired. THe table below represents mechanisms' relationships to security attributes.

Two important insights from the table and repeated by another example in the section on auditing are that 1.A SPECIFIC

CONTROL or MECHANISM CAN SATISFY MORE THAN ONE SECURITY ATTRIBUTE, and, 2.WHEN ONE SECURITY ATTRIBUTE IS CONTROLLED, OTHER OBJECTIVES MAY BE ATTAINED. For instance, encryption of the message text provides confidentiality. Encryption of a hash/control field (MAC) of the message text provides for data integrity. Encryption can provide authentication information. Encryption can be used for any or all of these security attributes. In addition, its use can result indirectly in non-repudiation as a result of the crypto key process where the possession of a particular key infers identity.

SECURITY MECHANISM/ATTRIBUTE ARCHITECTURE



Legend: Mechanisms that can provide a particular security attribute are indicated by X (ISO/IS 7498-2 p.7)

The descriptions in the two previous sections, "CONTROLS or SECURITY MECHANISMS" and "SECURITY ATTRIBUTES PROVIDED BY THE MECHANISMS," succinctly describe when matched by the reader why a particular mechanism can provide a particular attribute. For example:

ATTRIBUTE authentication:..... Peer entity provides verification that the party communicating with the receiver is the one that it claims to be.

MECHANISM authentication exchange: information exchange to ensure the validity of a claimed identity. Some examples are passwords, cryptographic techniques, and other characteristics and/or possessions of the entity being authenticated.

The table does not explicitly convey how the security mechanisms are to be used or their placement.

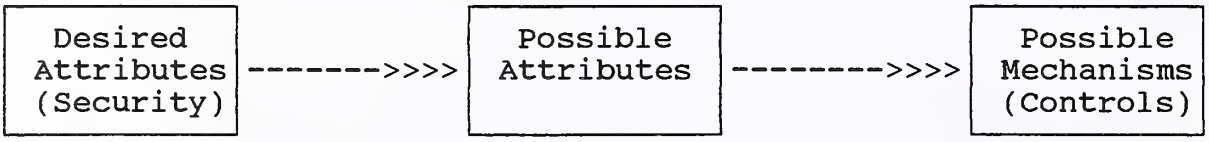
4. Planning the system:

The entity's system designers, guided by management, can develop an overall plan to provide a secure telecommunications environment. The emphasis is on the security potential that can be provided given the availability of technology (mechanisms).

There are two major strategies for designing system security. One is to design a secure system and to then "open it up" by granting privileges. The second is to have an unsecured system and patch on controls as needed. Most experts assert that controls should be not be patched, but should be an integral part of the design.

Step 1 in planning the system is, given the entity's security policy, assess the possibilities for security. The two previous tables (OSI layers & placement of security attributes and security mechanism/attribute architecture) are examples of useful materials. These provide attribute/mechanism possibilities to comply with the entity's security policy.

a. PLANNING THE SYSTEM



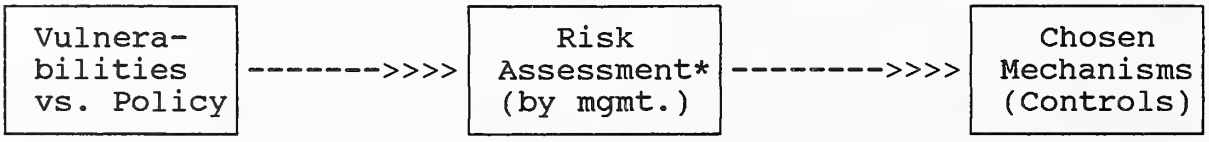
The actual design of a telecommunication working system can be undertaken. The tradeoffs of costs and benefits of various designs can be assessed. The traditional auditor framework separating potential problems into errors and irregularities can be applied in the telecommunication environment. One approach is to focus on threats or the overt attempts by a party to compromise the security of the system. The system design now can take place in an iterative process. Two decisions must be made:

1. the security mechanisms that should be available for any telecommunication message, and
2. the security mechanisms that should be used with a particular type of message[3].

For instance, consider a purchase cycle in which the entity's potential trading partners have proposed telecommunication transactions. The security attributes desired may increase as the message types progress from requests for quotations, through purchase orders, invoicing, and concludes with remittance advices including adjustments and telecommunication funds transfers (EFT).

For each message type above, the business decision is made about the needed level of security given the entity's overall security policy. That level of needed security determines attributes required and control mechanisms to be used.

b. DESIGN OF THE SYSTEM FOR MESSAGE TYPE



* A Risk Assessment[4] within the Entity's Security Policy

---

[3]. A extremely large number of security variations are possible for this decision point. For instance, security needs may vary not only by transaction type, but may vary also with different trading partners.

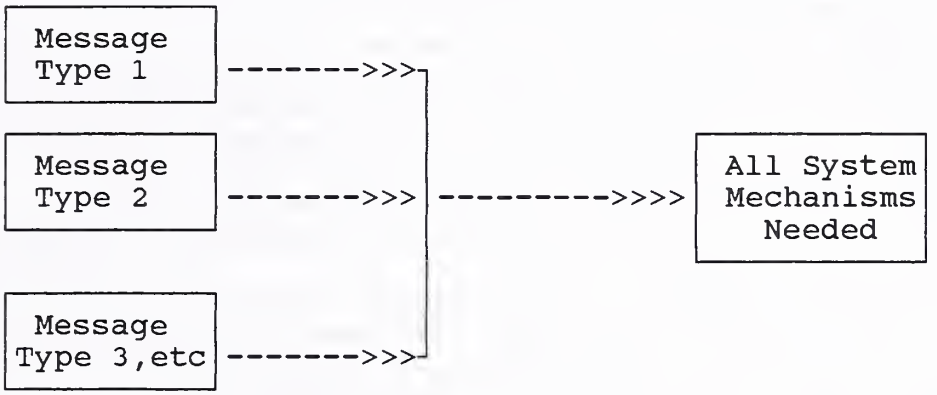
[4]. For a detailed example of EDI risk assessment see Burns, Mar, and Sorkin, Understanding and Auditing EDI and Open Network Controls, Bank Administration Institute and Institute of Internal Auditors. The methodolgy used conforms with GUIDELINE FOR AUTOMATIC DATA PROCESSING RISK ANALYSIS,

encompasses the assurance that the desired security will be provided (the Business Decision).

A quality assurance audit should take place at this point to ensure that the design complies with the entity's security policy. A risk assessment methodology must be used by management to determine rationally the specific mechanisms chosen. Controls should be chosen with an assessment of the risks and costs of controls. The design audit would include an evaluation of compliance with policy.

The mechanisms needed for each message type must now be reviewed to insure that these mechanisms will be or are now available. If the security mechanisms are not available,
1. they should either be made available,
2. or other mechanisms that can provide the same security attributes be used,
3. or a risk assessment be documented to justify not addressing that particular vulnerability(ies).

c. DESIGN OF THE OVERALL SYSTEM FOR NEEDED MECHANISMS



The iteration between mechanisms desired for each message type and the mechanisms made available is a management process. It continues within the framework of management risk assessment until those responsible approve the control design of the system. Once the system is designed, tested and implemented, the system's operations lend themselves to the traditional operational audit for each type of transaction.

FIPS PUB 65; and GUIDELINES FOR AUTOMATIC DATA PROCESSING PHYSICAL SECURITY AND RISK MANAGEMENT, FIPS PUB 31.

5. Terminology re Authentication and Nonrepudiation[5]:

The following is to provide a basis for understanding section 4 which describes some nonrepudiation scenarios.

Access control: This attribute is designed to prevent the use of the system by unauthorized personnel or in an unauthorized manner.

Assurance: This is meant to be the mitigator of risk. Assurance should be the result of risk assessment because it is attained at a cost. It may be a discrete cost such as using a particular smartcard technology in place of currently used password/PIN access. It can be a variable cost such as increasing the frequency of changing passwords or crypto keys. No attempt is made here to proscribe assurance because its cost should only be defended on the benefit basis associated with the mitigation of threats associated with the appropriate set of vulnerabilities. This is a business decision unless a particular assurance process is required by laws, regulations, or contracts.

Authentication: When trading partners are operating in an environment of "complete" trust with each other, the message receiver is concerned that the message sender can only be a particular message sender[6].

Digital signature: In a technical sense, this is a mechanism that both identifies the sender and can be used to provide a nonrepudiable proof of the sender, if needed, to third parties. This process is generally considered to use asymmetric (public key) cryptographic algorithms use two keys for cryptographic processing; a public key and a private key which are related in that one will be used for the encryption, the other key will be used to decrypt the message.
One key is published for use by any organization that wishes to communicate with the owner of the keys. The second key is kept secret by the owner and will reveal the information hidden by the use of the public key. Since public key algorithms are computationally expensive, this technique is most often used for short messages such as a MAC, such as the interchange of secret session keys which are used in some symmetric algorithm to encrypt and decrypt the data interchanged.

Encryption / the MAC: Authentication is a process of creating, transmitting and then verifying a security assurance field for a message. This assurance field must be some irreversible combination of all parts of the message that need protection and some secret data in such a way that the intended recipient can verify the authority of the originator to send the message and the authenticity (integrity) of the message received. Note that an authenticated message is transmitted as plaintext

---

[5]. The emphasis here is on sender authentication and nonrepudiation. Receiver authentication and nonrepudiation can be achieved through the response or conversational environment that occurs with EDI. Acknowledgements of protected EDI interchanges are needed to report the success or failure of the security process. The Functional Acknowledgment (transaction set 997) should be used to report failure of the security process. To ensure that it is from the actual receiver, the 997 itself should be authenticated.
The sending of a 997 without errors reported, or the sending of an Application Advice (transaction set 824) or a specific application response, is an implicit acknowledgement that the security process was successful.

[6]. If there is concern for a non-bonafide receiver, typically protection is provided by message encipherment.

(unless encrypted) and confidentiality is not provided. In situations where the hiding of message content is prohibited or not desired, authentication may be used.

Identification: This attribute is crucial because it is the basis for achieving all other security attributes. The term is used here in the simple sense of recognition of a participant. EDI identification is based upon many mechanisms such as passwords, PINs, smartcards, dedicated access transmission lines, third party logs, and cryptographic processes.

Nonrepudiation: When trading partners are operating in an environment of "partial" trust with each other, the message receiver is concerned that the message sender can only be a particular message sender and there is a "proof" available that will provide evidence to a third party that the sender a did in fact send that message.

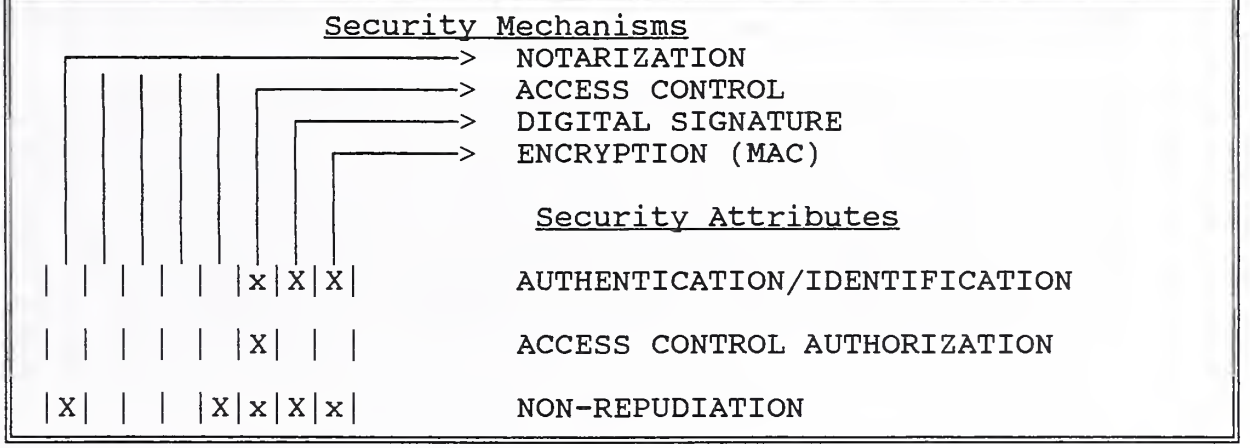
Notarization: This is the process of ensuring that additional evidence is available from a third party concerning the message sender and message contents.

6. Authentication and Nonrepudiation:

Elegance and text length limitations preclude a thorough discussion of these two security attributes from a business and technical standpoint. Message security is also not addressed from the ISO/OSI context, especially that may be provided by X400 and X500 implementations.

To provide a discussion framework, an abbreviated architecture table that was presented in Section 3 is below. The large X's are from ISO/IS 7498-2. The small x's have been added by the author. It is asserted that access control can provide authentication and, with the involvement of a third party, nonrepudiation. It is also asserted that the use of MACs with a third party can also provide nonrepudiation.

SECURITY MECHANISM/ATTRIBUTE ARCHITECTURE



THIRD PARTIES[7]

Third parties are defined in this section as entities that are independent of the trading partners and still can be participants in the EDI messaging environment. They can be:

IN-LINE: they handle the EDI message and, if they tamper with the message, it will not be detected.

ON-LINE: they handle the EDI message and, if they tamper with the message, it will be detected.

OFF-LINE: they do not handle the EDI message.

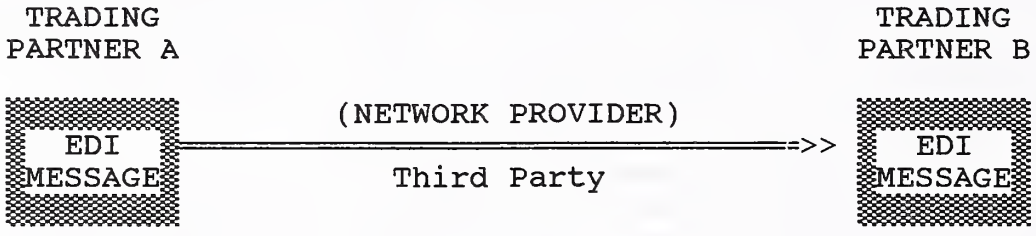
THIRD PARTY GOLDEN RULE

Third parties can provide for authentication if they do not keep logs (records). They can provide for nonrepudiation if they keep logs.

ACCESS CONTROL

EDI interchange requires three independent entities. They are portrayed below as trading partner A (EDI message sender), the network provider, and trading partner B (EDI message receiver). For the security attributes of authentication, it is assumed that the network provider functions as an on-line entity because it is trusted not to tamper with the message contents, or that, if not trusted, can not tamper because the trading partners are using MACing for content protection. This scenario is not an integral part of the X12 standards with the exception that the network provider can use X12 to communicate billing information.

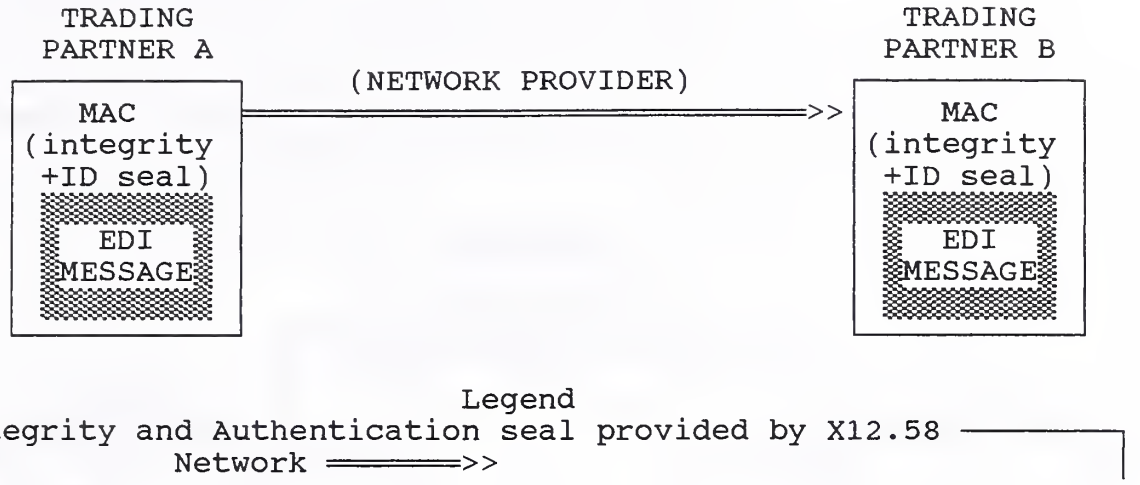
a. THE THREE EDI ENTITIES, TRUSTED NETWORK



Legend: Network ═══>>

[7]. Third parties can take many forms. They can be a VAN, a message handler such as a financial institution that forwards a remittance advice, a security facility that of a trading partner that is administratively independent, or a "black box" installed after the translator that is sealed and controlled by an external party such as an equipment supplier. The crucial feature is that they be independent of the trading partners' applications and not act in a biased manner toward either trading partner.

b. THE MAC, DISTRUST OF NETWORK PROVIDER



As portrayed, trading partner A may access the network by a password, smartcard, dedicated machine, or a dedicated secure wire. The network may then access B by any of the described means. B can then infer that A is who it is purported to be in the message. Nonrepudability is available in the form of testimony if the network provider keeps records of the message transfer. Only nonrepudability of the event[8] (a message) is available if the network does not record the message contents.

ENCRYPTION, THE MAC

ANSI/ASC X12.42 and X12.58 can be used for nonrepudiation with two basic scenarios:

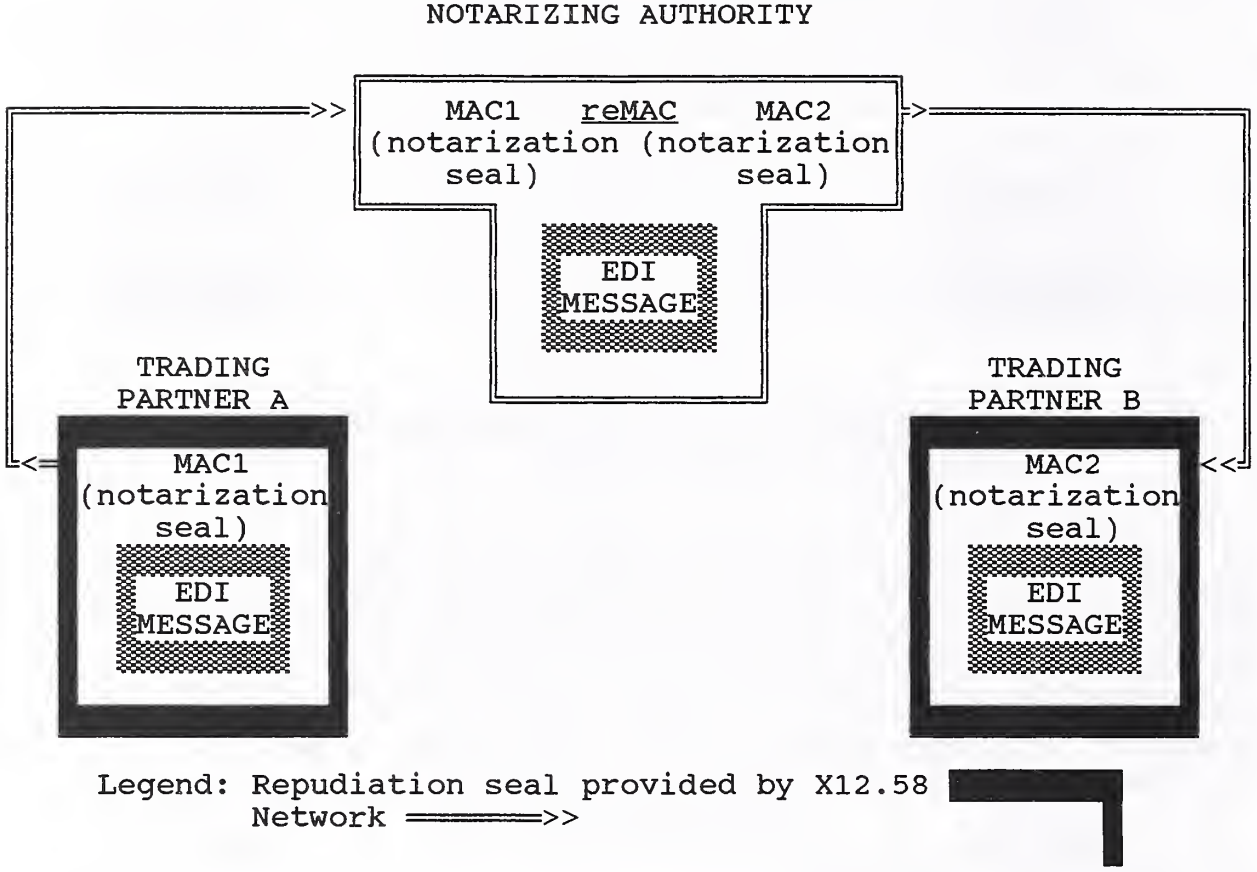
the third party is used to provide nonrepudiation and is "trusted" not to compromise message authenticity, integrity, and if desired confidentiality.

the third party is used to provide nonrepudiation and is not "trusted" regarding message authenticity, integrity, and if desired confidentiality.

Figure c represents a totally trusted third or notarizing party participation. The sender and receiver have unique and secret individual cryptographic keying relationships with the notarizing party. The sender MACs (MAC1) the EDI message at either the functional group or transaction set level by means of its agreed to keying relationship with the notarizing authority and transmits to the notarizing authority. The notarizing authority checks the message and its associated MAC (MAC1) for agreement to ensure that it is authentic (came from the receiver).

[8]. In the real world, the network provider does keep event records to substantiate billing activities for rendered services.

c. NONREPUDIATION AND FULLY TRUSTED NOTARIZER



The notarizing authority then reMACs (MAC2) the message by means of its agreed to keying relationship with the receiver and transmits to the receiver. The receiver checks the message and its associated MAC (MAC2) for agreement to ensure that it is authentic (came from the sender). The receiver therefore concludes that the EDI message came originally from the sender because of the trust in the notarizing authority and the strong assurance associated with the use of X12.48 and X12.58 security methodology.

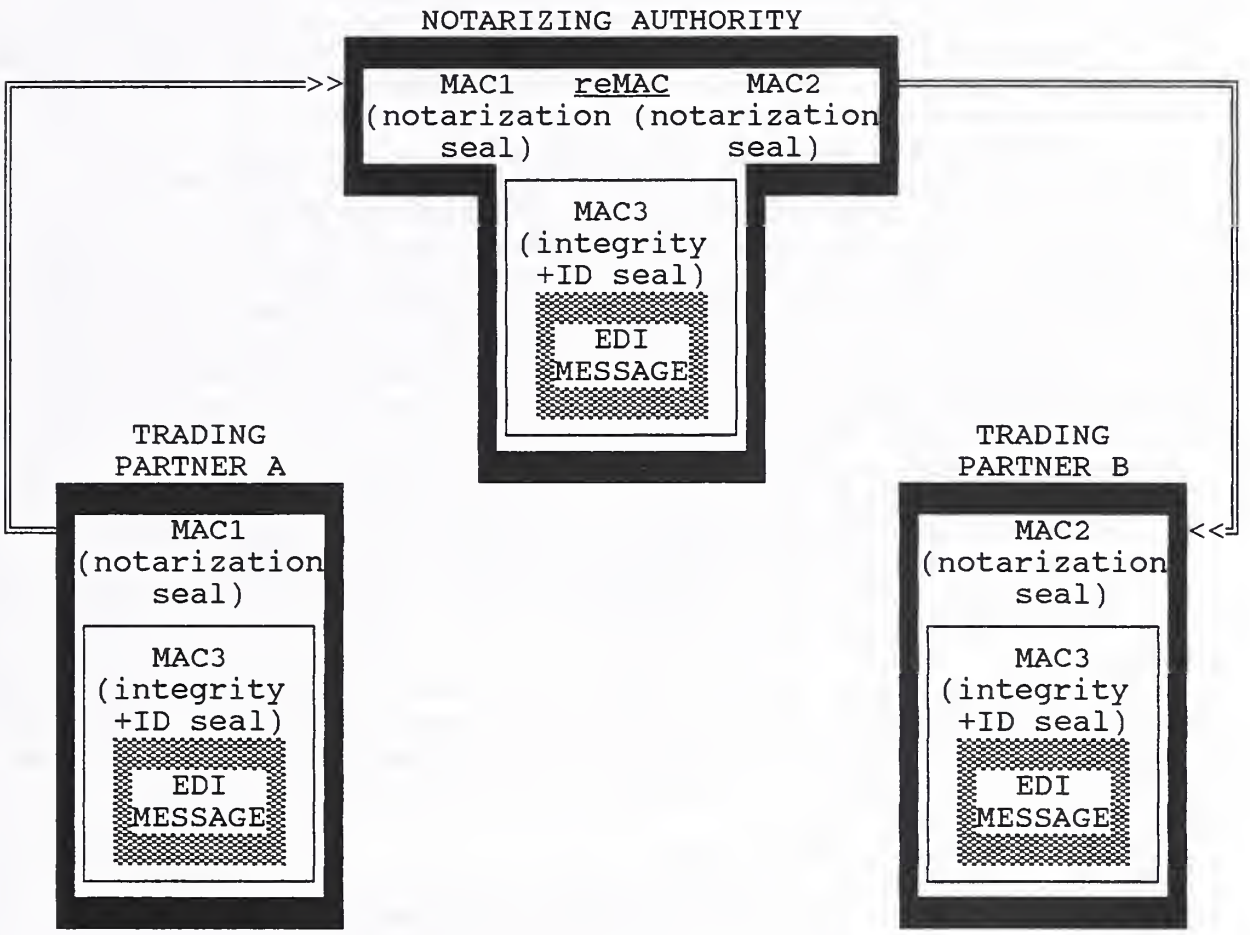
If the sender should attempt to repudiate the message, the receiver can reconstruct the cryptographic process with the aid and testimony of the notarizing authority and it can be therefore concluded that particular EDI message could only have been generated by the sender who is attempting repudiation.

Of course, records will have to be maintained by both the receiver and the notarizing authority. In particular, the receiver will have to be able to reconstruct the originally received message, either the notarizing authority or the receiver will have to be able to reconstruct MAC2, the notarizing authority will have to maintain a record of the keying relationships that existed to create the particular MAC1 and MAC2, and the trusted notarizing authority should have a record of the particular MAC1 used for the message that is not disclosed until testimony is required.

This is an in-line scenario. The third party can tamper with the message contents and not be detected by B.

Figure d represents the situation when the sender and receiver do not completely trust the third or notarizing party in the sense that they wish to be assured that the notarizing authority has not tampered with message contents. In this case, the sender, notarizing authority and receiver utilize the nonrepudiation process (MAC1 and MAC2) at the functional group level.

d. NONREPUDIATION AND PARTIALLY TRUSTED NOTARIZER



To detect possible tampering by the notarizing authority with message contents, the sender and receiver proceed to do a pairwise MACing (MAC3) at the transaction set level. And, of course, appropriate records will be kept by all three parties. This is an on-line scenario because message contents tampering by the third party can be detected by B.

DIGITAL SIGNATURE

At this time there is not the provision to use public key digital signature technology within the context of X12 standardized EDI formats. It is anticipated that future X12

standards will accommodate digital signatures, tokens, and other security assurances. Theoretically, none of these are required to achieve authenticity and nonrepudiation. On a more practical level, it is anticipated that these should be accommodated to provide different cost and technical trade-offs, and also to accommodate entities that have other rationales to use something besides DEA based MACing.

7. Auditors and Auditing:

Why the auditors? Although this is covered in detail elsewhere, the auditors ensure that the system designers have answered entity policy questions regarding the compliance with

a. Government Auditing Standards, 1988 Revision, (The Yellow Book), United States General Accounting Office by the Comptroller General of the United States

b. AICPA Professional Standards, Volumes 1 and 2, as of June 1, 1992, American Institute of Certified Public Accountants, especially SAS 55, AICPA's Internal Control Standard.

Typical concerns management attempts to address in designing the internal control structure are (a) to provide reliable data, (b) to safeguard assets and records, (c) to promote operational efficiency, (d) to encourage adherence to prescribed policies and (e) to comply with laws and regulations. Reasonable assurance should be provided by the internal control structure so that the following seven detailed objectives are met:

(1) recorded transactions are valid (validity)
(2) transactions are properly authorized (authorization)
(3) existing transactions are recorded (completeness)
(4) transactions are properly valued (valuation)
(5) transactions are properly classified (classification)
(6) transactions are recorded at the proper time (timing or cut-off)
(7) transactions are properly included in subsidiary records and correctly summarized (posting and summarization).

Let us assume that on the basis of experience, we know that the following 12 types of controls can be used in a claims reimbursement process:

1.Unissued checks controlled
2.Check signature plate physically secured
3.Checks reviewed and approved prior to issue
4.Checks issued agreed to cash request report
5.Claimant statement independently reviewed and reconciled
6.Accounts payable account in general ledger reconciled to accounts payable ledger
7.Checks issued balanced to daily payments journal
8.Monthly bank reconciliation prepared
9.Independent review of monthly bank reconciliation
10.Claimant payable account in general ledger reconciled to accounts payable ledger
11.General ledger account code checked on request for payment
12.Monthly payments journal reviewed for reasonableness

To summarize the relationship of the use of these 12 controls to attain these seven specific objectives, the following table is presented. Two important insights from the table are that 1. a specific control or mechanism can satisfy more than one objective, and, 2. When one objective is controlled, other

objectives may be attained.

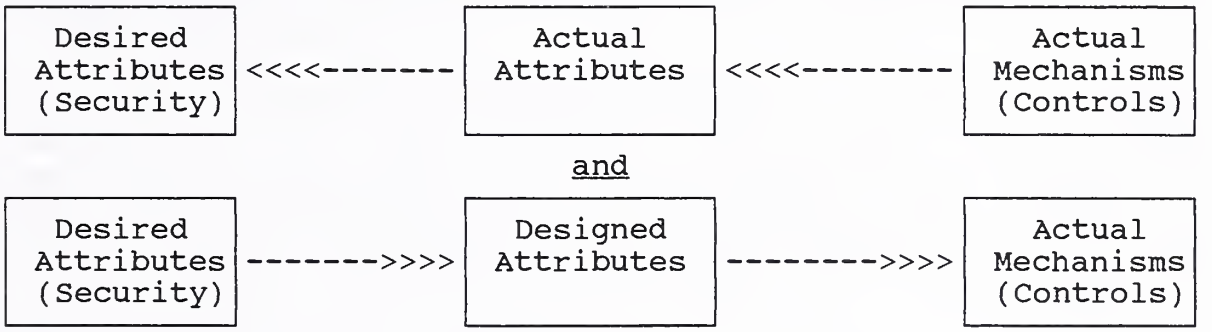
RELATIONSHIP OF REIMBURSEMENT OBJECTIVES AND CONTROLS USED	
REIMBURSEMENT CRITERIA	APPLICABLE SPECIFIC CONTROLS
1.EXISTENCE- Validity	1,2,3,4,5
2.EXISTENCE- Completeness	5,6,7,8,9
3.EXISTENCE- Recording	7,8,9,10
4.VALUATION- Accuracy	5 7,8,9,10
5.CUT-OFF	5 7,8,9
6.CLASSIFICATION	10,11,12
7.MECHANICAL ACCURACY	8,9,10,11,12

Notice that control activities 8 and 9 associated with bank statements and reconciliations appear especially powerful in a claims reimbursement area. These two activities are controls that can help achieve five out of the seven reimbursement objectives.

Audit, in the above development process, should review and assess the control design. Recommendations should be made if the design does not comply with the entity's security policy. If an audit was not performed until the system is implemented and the design did not provide for adequate security, then it may be too costly and nonproductive to redesign the system. If the system is inadequate, audit fieldwork of its operation would have little meaning because the system is not in compliance with policy.

After the system is implemented, the compliance audit of operations will take place periodically.

d. AUDIT OF THE SYSTEM (operations)



The auditor is involved in reviewing and evaluating the three planning and design phases:

a. PLANNING THE SYSTEM,

b. DESIGN OF THE SYSTEM FOR MESSAGE TYPE, and

c. DESIGN OF THE OVERALL SYSTEM FOR NEEDED MECHANISMS.

The auditor then is concerned with the implementation in the next step,

d. AUDIT OF THE SYSTEM (operations).

The Yellow Book requires the auditor to know a client 's internal control structure well enough to plan the work and to select procedures appropriate to the objectives of the engagement. To gain knowledge of the internal control structure, an auditor studies its elements: the control environment, the

accounting system, and control procedures. This section outlines these elements, the objectives of internal control, and the procedures auditors might use to satisfy auditing standards.

Other things being equal, the less control risk (stronger the internal control structure) there is, the less audit risk faced by the auditor. The elements of the internal control structure should be cost beneficial. Typical concerns management attempts to address in designing the internal control structure are to provide reliable data, to safeguard assets and records, to promote operational efficiency, to encourage adherence to prescribed policies, and to comply with laws and regulations.

Auditors are primarily concerned with controls that provide reliable data and safeguard assets and records because these directly affect the financial statements and assertions that management makes about balances appearing therein.

There are four assumptions (concepts) which underlie the study of an internal control structure. It is management's responsibility to establish and maintain the internal control structure. Reasonable but not absolute assurance should be provided because an ideal system cannot be justified on a cost/benefit basis. Even the ideal internal control structure has inherent limitations because of employee carelessness, lack of understanding, or management override. As a result, control risk is always positive. Control concepts, objectives, and the methodology of assessment are applicable in either a manual or computerized structure. Reasonable assurance should be provided by the internal control structure of each transaction cycle that the previously discussed seven detailed objectives are met.

There is relatively few new audit concerns with EDI at a high level. Even the audit of cryptographic environments have been described fully. Emphasis should placed on keeping abreast of new laws and regulations. The importance of balancing EDI activities with network providers is trivial to mention, and likewise the concern for adequate records. The loss of controlled and prenumbered paper stock is supplanted by concerns for controls over counters, date/time mechanisms, and authorization and internal access controls.

It does not appear to make much sense to be concerned about EDI transmission security if internal controls at the business application level are inadequate.

WHAT PRICE DATA SECURITY?

by
John L. Stelzer
Sterling Software, Inc.

The time- and content-criticality of business information being moved via Electronic Data Interchange (EDI) have been steadily increasing over the past several years. Accordingly, the attention paid to audits, controls, and security has gradually heightened. Over time, as organizations begin to move business transactions that have significant dollar or strategic implications, the need for reliable, secure, and traceable information control methodologies will become nothing short of a necessity.

Today, as organizations have begun to seriously consider the potential impact of business data that might get lost, duplicated, compromised, modified, delayed, etc., they have come to the realization that, to some extent, paranoia is just good thinking. Unfortunately, there is a tendency–when considering the risk and implications of potential EDI calamities–to go overboard with the paranoia. The result can be security, audit, and control measures whose cost far exceeds the value of the business information being protected. In an effort to identify cost-effective and, at the same time, fully adequate safety mechanisms, this writing will seek to identify: security, audit, and control objectives; key points of concern; methodologies for addressing these concerns; and the application and relative effectiveness and cost of each approach. It is intended to be an overview of the available mechanisms rather than an in-depth discussion of the functions of each.

Since this topical area is extremely broad, focus will be restricted to the one area that is usually least well-understood by EDI participants, that is, the movement of EDI data between the EDI systems of the sender and the receiver. Clearly, attention must also be paid by EDI users to security, audit, and control concerns for data as it is handled in the EDI systems themselves and, of course, in the internal business systems of each participant. However, data movement between trading partners' EDI systems can encounter numerous variables and raise unique challenges for those concerned with security and data tracking. In addition, the potential insertion of one (or more) EDI network service provider(s) into the equation adds new and interesting questions, additional potential for failure, and, ironically, opportunities for improved control.

SECURITY, AUDIT, AND CONTROL OBJECTIVES

In general, there are six key security and control objectives when moving data electronically between trading partners:

- **Content Integrity** requires that the recipient be able to verify that the document's content has not been altered.

- **Sequence Integrity** requires that the recipient have the ability to identify skipped, duplicated, or out-of-sequence documents.

- **Content Confidentiality** requires that the originator and recipient be able to ensure that the document's contents are not subjected to unauthorized disclosure.

- **Originator Authentication** requires that the recipient be able to verify that the purported originator is, indeed, the actual originator.

- **Recipient Authentication** requires that the originator and recipient be able to verify that the document has actually been received by only the intended recipient.

- **Timely Delivery** requires the use of a mechanism that can verify that the document was delivered by a given time.

The criticality of each of these objectives may vary by: document type, document content-criticality, document time-criticality, and, perhaps, trading partner. As such, the extent to which an organization is willing to go and the level of expense the organization is willing to incur to assure that these objectives are consistently met will change in direct relation to the importance of these variables. In general, the degree to which an organization can suffer harm if a given objective is not met should be the primary driver that determines the level of investment that the organization will make to ensure that the objective is met.

POINTS OF CONCERN

There are four external data transfer environments to which these six objectives must be applied:

- **Direct (Point-To-Point):** EDI transactions move directly from the EDI system of the originator to the system of the recipient (without passing through an EDI network service provider) (See Figure 1).

- **Network:** Both trading partners exchange EDI transactions with each other via the same EDI network service provider (See Figure 2).

- **"Public" Interconnect:** The originator and the recipient each use a different commercial EDI network service provider. These networks exchange the data via network interconnect links on behalf of their respective customers (See Figure 3).

- **"Private" Interconnect:** One of the participants in the trading pair chooses not to use a commercial EDI network service provider, while the other partner does. The resulting link between the network and the "off-network" trading partner is considered a "private" interconnect because it requires the commercial EDI network to interconnect with a company that is not traditionally in the business of commercially providing EDI network services (See Figure 4).

Each of these four data transfer environments has its own potential points of concern when considering the security and control objectives. This section will provide a brief

overview of each before going on to look at each of the security and control objectives as they apply to each of these environments. At that time, mechanisms for achieving these objectives in each of the environments will be discussed.

Direct (Point-To-Point):

The direct link may or may not utilize a signon/password sequence. The originator of the document may or may not initiate the call to the recipient, and data may be sent and/or received during the session. Most current direct links are send only, originator-initiated sessions that do not use much in the way of a signon/password sequence. Where such a signon sequence *is* used, its form, procedures, and effectiveness typically vary widely from link to link.

Strengths:

It can generally be agreed that the fewer parties handling the data, the fewer points of vulnerability. Certainly, the absence of interim data "handlers" bodes well for timely delivery. There are, however, several possible areas of concern.

Weaknesses:

While it is generally assumed that in a direct link no one else has handled or seen the data, it is certainly possible for someone to intercept and view, modify, duplicate, or discard the document. In a direct link, the functional acknowledgment is usually the only control response that is provided by the recipient to the originator. It is possible–and, in many cases, reasonably easy–for the impostor to return a functional acknowledgment as if it had come from the intended recipient. If the intent of the impostor is to modify the document and pass it on to the originally intended recipient, the impostor could successfully transmit the modified document to the recipient's system fairly easily–given the general lack of inbound access security present in most of today's direct links–and still avoid detection because of the delay–due to the lack of a mechanism to report data receipt time in the EDI acknowledgment . This potential for such a difficult to detect intrusion by an impostor puts in some jeopardy most of the security and control objectives as they apply to direct links.

Network:

In most cases, networks require use of a special signon/password procedure to validate the identity of the party attempting to access the network. In many cases, networks maintain network customer profiles that define various security parameters (e.g. signon id and password, special network command ids and passwords, valid EDI sender and receiver ids, etc.). Some networks maintain communications profiles to further assist in validating a given party attempting to access the network.

Virtually all networks allow send only, receive only, or send and receive sessions to be conducted. Some networks provide special processing services for their customers such as format translation, envelope conversion, syntax checking, carbon copying, etc.

Most network users request that transactions destined for them be held in network storage areas known as mailboxes until the user can pro-actively access the network and retrieve the data. Many networks offer a variety of options for delivering the data to the end recipient including various forms of network-initiated outdial and translation to and transmission in other media forms such as facsimile, mail, or electronic mail.

Some networks provide data status control reporting to originators and recipients on their network. Such reporting typically informs the originator which transactions were received by the network, whether each transaction could be processed by the network, whether and when each transaction was posted to the recipient's mailbox, and when each transaction was retrieved from the mailbox by the recipient. Network control reporting for recipients typically identifies the identity of transactions posted to the recipient's mailbox, the date and time of posting, and the date and time when the network shows that each transaction was retrieved by the recipient.

Strengths:

The addition of access security measures and customer profiles by networks makes the job of the impostor much more difficult. Such additional security measures contribute to the certainty of originator and recipient authentication. The use of control reporting to the originator and the recipient also lessens the likelihood of undetected oddities and contributes to sequence integrity.

Weaknesses:

Handling of the transaction by an additional party (namely the network) *can* introduce the possibility of content integrity problems and content confidentiality breeches. In addition, network processing can create problems with the timely delivery of data.

"Public" Interconnect:

Most interconnects between commercial EDI network providers utilize proprietary signons, the Interconnect Mailbag, or X.400/X.435. A few still utilize the older sender-initiated "dial, dump, and pray" method of interconnect. Most interconnects are sender-initiated send-only sessions, although a few variations remain.

Many networks maintain interconnect profiles (similar to their normal customer profiles) to identify valid ids, passwords, and, in some cases, the communications behavior of interconnecting networks. A very small handful of networks also maintain sender origin verification tables to confirm that a transaction from a given originator should have been received via a particular interconnect link. This further complicates the life of the impostor.

The proper destination of an outbound interconnect transaction is determined by reading an "off-network" partner profile that identifies that recipient's interconnect network.

Strengths:

The recent addition of standardized logon mechanisms for interconnecting networks has significantly improved the security of public interconnects. Furthermore, the use of either the Interconnect Mailbag (and the accompanying Interconnect (receipt) Acknowledgment) or the X.400/X.435 control mechanisms has provided the necessary assurance that transactions were not lost or duplicated during the interconnect exchange. This contributes significantly to sequence integrity. The use of sender origin verification and off-network partner profiles contributes to originator and recipient authentication respectively.

Weaknesses:

Some networks elect to queue outbound transactions destined for interconnects until some minimum volume has been met. Others hold inbound transactions received from interconnects for some time before processing the data. Still others assign interconnect transactions (outbound or inbound) lower processing priorities than non-interconnect transactions. All of these practices have negative implications for timely delivery. Use of key control point dates and times can allow the originator and recipient to easily detect such practices, however.

"Private" Interconnect:

Most private interconnects do not use any extensive form of logon (either when going commercial-to-private or vice versa) because of the non-commercial element in the interconnect. Much like direct links, private interconnects are usually sender-initiated, send-only transaction transfers.

Strengths:

Because standardized logons are rarely used, private interconnects forfeit many of the same improved controls that the direct link must also forego. In particular, for data coming from the private interconnect side, the receiving network has no identification of the sending interconnect party other than the sender interchange EDI id in the data. The receiving commercial network cannot, therefore, authenticate the originator or perform sender origin verification.

The commercial-to-private portion of the interconnect remains reliable since the off-network profile will identify the unique phone number of the recipient to be called and, therefore, improve the reliability of recipient authentication.

Weaknesses:

As with public interconnects, timely delivery can be negatively impacted if the commercial network chooses to treat interconnect data differently than non-interconnect data. Once again, monitoring of control point dates and times can allow the originator to detect such delays.

SECURITY, AUDIT, AND CONTROL METHODOLOGIES

There is a wide variety of useful tools for addressing the six security and control objectives. They vary in effectiveness, required effort, and cost. Again, those considering security and control alternatives should carefully balance the dollar and strategic value of the threat with the cost of the protection. This section discusses the control approaches that are available or could be instituted to satisfy each of the six security and control objectives at varying levels of effectiveness and cost.

I. Content Integrity

Data Authentication: The most conclusive assurance of non-alteration of data comes from the use of data authentication. Available in several forms and using any one of a variety of different algorithms, data authentication allows the recipient to be sure that no portion of the authenticated data has been modified in any way. Authentication is also one of the more expensive solutions since software must usually be purchased by each of the trading partners to perform the authentication process. In addition, because there is no universal authentication standard, it is likely that a given organization will have to purchase several different authentication software packages (or one robust package capable of handling many different authentication approaches) in order to accommodate variations from partner to partner.

Alternatives: A somewhat less conclusive (but significantly less costly) approach for verifying content integrity incorporates a combination of tools available to most network users. To best illustrate how these tools work, consider the ways that data might be altered as it moves from the originator through the network(s) to the recipient.

I.1 Alteration From Communication Errors: When EDI data segments are "wrapped" (i.e. spanning individual communication units such as records, blocks, packets, etc.), communication errors can be detected by CRC protocol checking and/or syntax variations. Most communication errors (that are not caught by CRC protocol checking) involve the loss or duplication of entire communication units. If data segments are allowed to span these individual communication units (the common practice in most EDI transmissions, today), such loss or duplication of communication units results in syntactically invalid EDI data which can then be detected by the recipient's EDI translation software.

I.2 Manual Alteration Of Data Within The Network: Using data mirroring, internal cross-checks, and properly secured and controlled access to data within the network system, the possibility of manual manipulation of data while in the EDI network system can be reduced to a non-issue. Concerned originators and recipients should request to see results of technical audits performed on their network service provider(s) operations.

I.3 Manual Alteration Of Data During Data Communication: There are three possible scenarios here. Data is captured and manipulated: (1) between the originator and the originator's network (before network receipt of the data), (2)

between two interconnecting networks (before receipt of the data by the recipient's network), or (3) between the recipient's network and the recipient. Each is addressed separately:

I.3.1 Capture Before the Originator's Network: For this capture and manipulation to occur successfully, the following must take place:

- ❑ An impostor must:
  - Capture the transactions without the originator or the network detecting any communication oddities (i.e. the impostor must be able to either mimic the network behavior to the originator or capture the data without session disturbance before the transactions reach the network). This also presumes that the impostor is aware of <u>when</u> the originator will access the network.
  - Manipulate the data without changing any character or segment counts
  - Use EDI interchange sender and receiver ids (if part of the manipulated data) that are considered valid by the originator's and recipient's network(s) and by the recipient
  - Know and use the originator's network signon and password sequence to log on to the network to insert the data back into the delivery process
  - Mimic the originator's communication behavior as specified in the network's customer profile (protocol, line speed, block size, and options such as: transparency, compression, etc.)
  - Manipulate the transactions and send them to the network in a time frame that is short enough to avoid raising suspicions on the part of the originator (e.g. detection of a wide disparity between the originator's communications transmit time and the network's reported receipt and processing time). Note: This implies that networks should start supplying date and time of file receipt in their control reporting.
  - (If sender and/or receiver ids and/or character or segment counts were changed) Retrieve or intercept the network's control reports that cite the "new" or manipulated ids and/or counts <u>before</u> the reports fall into the hands of the originator. Note: Since the absence of such a report would similarly cause the originator to become suspicious, the impostor would also have to modify the report and somehow get it into (a) the mailbox of the originator or (b) the hands of the originator during a subsequent network access by the originator.

I.3.2 Capture The Data During The Interconnect: For this capture and manipulation to occur, the following must happen:

- ❑ An impostor must:
  - Intercept the data during the interconnect communication session without the interconnecting networks detecting any communications oddities. This also presumes that the impostor is aware of <u>when</u> the originating network will initiate the interconnect.
  - Generate and return to the originator's network the appropriate receipt notification (With the Interconnect Mailbag, for instance, the impostor would have to return an Interconnect Acknowledgment using the proper

authorization and security ids and sequential mailbag control number, all within the previously established acknowledgment response time setup between the two interconnecting networks.)
- Successfully interconnect and transmit the manipulated data to the recipient's network (using the appropriate interconnect control enveloping, ids, control numbering, etc.)
- Intercept the receipt notification (e.g. Interconnect Acknowledgment) coming from the recipient's network before it is received by the originator's network

I.3.3 Capture The Data Between the Recipient's Network And The Recipient:

- ❑ The impostor must:
  - Capture the transactions without the recipient or the network detecting any communication oddities (i.e. the impostor must be able to intercept the data without session disturbance and before the transactions reach the recipient). This also presumes that the impostor is aware of <u>when</u> the recipient will access the network.
  - Manipulate the data without changing any character or segment counts
  - Manipulate the transactions and get them into the hands of the recipient in a time frame that is short enough to avoid raising suspicions on the part of the recipient (e.g. detection of a wide disparity between the recipient's actual reception time and the network's reported mailbox receipt time). Note: This assumes that the recipient's network supplies the date and time of data retrieval from the mailbox in their control reporting. Also note, that to get the data into the hands of the recipient, the impostor would have to (1) transmit the data into the network using a signon that was valid for the EDI sender id used in the interchange, intercept and appropriately modify the recipient's control reports to remove all references to the re-transmission of the modified data, and get the modified control report into the hands of the recipient or (2) insert the modified data into a recipient communication session with the network in such a way that neither the recipient nor the network could detect the communications interruption and the recipient would receive the extra data.
  - (If sender and/or receiver ids and/or character or segment counts were changed or the time to return the manipulated data to the recipient's mailbox took too long) Retrieve or intercept the network's control reports that cite the "new" or manipulated ids and/or counts or the impostor's pickup time <u>before</u> the reports fall into the hands of the recipient. Note: Since the absence of such a report would cause the recipient to become suspicious, the impostor would also have to modify the report and somehow get it into (a) the mailbox of the recipient or (b) the hands of the recipient during a subsequent network access by the recipient.

II. Sequence Integrity

EDI Control Numbering: By properly using envelope control numbering mechanisms available in the EDI standard, the recipient is able to identify: the sequence intended by a given originator, duplicated interchanges, and skipped interchanges. While the standard provides no specifics for how these numbers should be sequenced, an originator can provide the tools necessary for a recipient to verify sequence integrity by sequentially incrementing interchange control numbers on a per recipient basis. When this approach is followed, the recipient can identify skipped, duplicated, or out-of-sequence interchanges using the interchange control number. Note: It is recommended that the originator begin with a zero control number and only return to zero for a given recipient when the incremented control number naturally rolls over to zero again. Originators should not, for instance, reset their control numbers to zero at the beginning of each day, week, etc.

Additional control can be obtained by labeling functional group and transaction set control envelopes with control numbers that are hierarchically related to that of their "parent" envelope (For instance, the first functional group in interchange number 7984 might be numbered 798401. The first transaction set in that first functional group might be numbered 79840101, and the second transaction set in that same functional group might be 79840102, and so on.)

III. Content Confidentiality

Encryption: The most conclusive assurance of content confidentiality comes from the use of data encryption. Available in several forms and using any one of a variety of different hardware and/or software combinations and algorithms, data encryption allows the recipient to be certain that no portion of the data has been disclosed to unauthorized parties. Unfortunately, encryption is also one of the more expensive solutions. Hardware and/or software must usually be purchased by each of the trading partners to perform the encryption/decryption process. Because there is no universal encryption standard, it is likely that a given organization will have to purchase several different encryption packages (or one robust package capable of handling many different encryption approaches) in order to accommodate variations from partner to partner. In addition, encrypted data typically contains characters that can conflict with reserved communication characters used in several of the most widely implemented protocols. This conflict can be resolved if the protocol supports some form of transparency that allows the protocol software to disregard certain streams of characters. Unfortunately, not all protocol software packages have a transparent mode. Another solution is to change to a protocol whose characters do not conflict with the encrypted characters. This is not always practical nor economical. A third approach is to filter the data, thereby converting all characters to values outside the reserved protocol character set. Unfortunately, nearly all data filters tend to double the character count. For most network users, this can double their network traffic costs since most networks charge based on character volume.

Alternatives: A somewhat less conclusive (but significantly less costly) approach to assuring content confidentiality relies on many of the same impostor impediments discussed under content integrity. To best illustrate how these impediments deter the impostor, consider the ways that data might be viewed as it moves from the originator through the network(s) to the recipient.

I.1 Impostor Access To Data Within The Network: Using proper internal network security and controlled access to data within the network system, the possibility of unauthorized access to data while in the EDI network system can be reduced to a non-issue. Concerned originators and recipients should request to see results of technical audits performed on their network service provider(s) operations.

I.2 Access To Data During Data Communication: There are three possible scenarios here. Data is captured: (1) between the originator and the originator's network (before network receipt of the data), (2) between two interconnecting networks, or (3) between the recipient's network and the recipient. Each is addressed separately:

I.2.1 Capture Between the Originator and the Originator's Network: For this capture to occur successfully, the following must take place:

- ❑ The transactions must be captured without the originator or the network detecting any communication oddities (i.e. the impostor must be able to either mimic the network behavior to the originator or capture the data without session disturbance before the transactions reach the network). This also presumes that the impostor is aware of <u>when</u> the originator will access the network.

I.2.2 Capture The Data During The Interconnect: For this capture to occur, the following must happen:

- ❑ The transactions must be captured during the interconnect communication session without the interconnecting networks detecting any communications oddities. This also presumes that the impostor is aware of <u>when</u> the originating network will initiate the interconnect.

I.2.3 Capture The Data Between the Recipient's Network And The Recipient:

- ❑ The transactions must be captured without the recipient or the network detecting any communication oddities (i.e. the impostor must be able to intercept the data without session disturbance and before the transactions reach the recipient). This also presumes that the impostor is aware of <u>when</u> the recipient will access the network.

- ❑ The impostor could alternately access the recipient's mailbox to retrieve the transactions. To accomplish this without detection, the impostor would have to:

- Access the recipient's mailbox using network-assigned security ids and passwords
- Successfully mimic the communications behavior of the recipient consistent with the characteristics stored in the network's customer profile for the recipient
- Retrieve the mailbox contents using the proper network commands (including any additional ids and passwords) in the proper session sequence as required by the network
- Retrieve or intercept the network's control reports that cite the pickup time of the impostor before the reports fall into the hands of the recipient. Note: Since the absence of such a report would cause the recipient to become suspicious, the impostor would also have to modify the report and somehow get it into (a) the mailbox of the recipient or (b) the hands of the recipient during a subsequent network access by the recipient.
- Cause a copy of the accessed transactions to be placed back into the recipient's mailbox to avoid suspicion caused by missing data. This could be done by causing the network to restore the data (although use of control reports reminding the requester of all restore activity would cause the impostor to have to intercept the restore control report, as well).

❑ Alternately, the impostor could transmit the data into the network to be loaded to the recipient's mailbox. To do this, the impostor would have to:
- Know and use the originator's network signon and password sequence to log on to the network to insert the data back into the delivery process
- Mimic the originator's communication behavior as specified in the network's customer profile
- Retrieve or intercept the network's control reports that cite the re-transmission of the data before the reports fall into the hands of the originator
- Intercept, modify, and return to the originator any other control reports which would show the transmit and new pickup time of the re-transmitted data.

IV. Originator Authentication

Authentication: Once again, authentication provides one of the most conclusive mechanisms for authenticating originator identity. As has been previously discussed, however, authentication is not without its associated costs.

Alternatives: A less expensive, and yet effective, approach involves the combined use of many of the impostor barriers discussed earlier. Given the array of network controls and security devices in place in most conscientious network providers, the barriers to the would-be impostor are sufficiently high to cause a lack of interest in all but the most valuable transactions. Signon ids and passwords; communications profiles; network session scripts and commands (with *their* special ids and passwords); network originator notification of network data receipt (via control reports); network matching of the EDI sender id with the signon id and

password used; trading pair definitions; etc. all serve to thwart efforts of impostors to send data on behalf of someone else. As such, they provide reasonable assurance that data received by a network through these security filters is, indeed, from the purported originator.

Where interconnects are involved, the use of interconnect controls such as the Interconnect Mailbag or X.400/X435 in conjunction with mechanisms to verify the correct origin of a particular originator (via a given network) can similarly provide reasonable originator authentication. These mechanisms position the network to be able to provide the recipient with reasonable originator authentication. This notice of authentication is implicit in that the data would not have been posted to the recipient's mailbox if it had not passed all of the security filters. Therefore, presence in the recipient's mailbox implies originator authentication. For those recipients who do not wish to have their network suspend data for <u>any</u> security reasons, the recipient's network could alert the recipient about an originator that could not be authenticated (using normal network control reports that flagged entries that referred to suspect data).

V. Recipient Authentication

Receipt Assurance: Proof that the intended recipient actually received the data can be accomplished using any one of several receipt acknowledgment mechanisms available (e.g. Functional Acknowledgment, TA1, an application/business acknowledgment, etc.). While any one of these could be returned by an impostor to mimic the receipt acknowledgment actions of the intended recipient, use of unique authorization and security ids in the interchange header create additional barriers to entry for the would-be impostor (once again, making all but the most rewarding documents unattractive). It should be noted that if an interim data handler (i.e. EDI network, X.400/X.435 MHS, etc.) does not first receive a pro-active acknowledgment of receipt from the end recipient–in a form other than at a communications protocol level–then any claim made by that intermediate handler that the data was actually received by the end recipient is inconclusive.

There are several additional tools available through most EDI networks that allow the end recipient to have a high level of certainty that they have received all of the data intended for them. These tools take the form of various network control reports indicating the transactions that were posted to the recipient's mailbox, the date and time of posting, and the date and time of retrieval from the mailbox by the end recipient. These same control reports allow the end recipient to detect oddities with respect to unauthorized access to data (e.g. a pickup date and time listed for transactions that were never accessed by the recipient).

As with the impact of network security mechanisms on sender authentication, signon ids and passwords; communications profiles; network session scripts and commands (with *their* special ids and passwords); network recipient notification (via control reports) of network data posting; network matching of the EDI receiver id with the appropriate recipient mailbox; trading pair definitions; etc. all serve to

assure the originator of the correct posting of the data and the receipt of that data by the intended recipient. As mentioned under Confidentiality, they also serve to thwart efforts of impostors to retrieve data on behalf of the intended recipient.

VI. Timely Delivery

Time-Based Data Tracking: As the time-criticality of EDI data movement grows in importance, tools that have long been taken for granted will increase in value. Key milestones in the life cycle of EDI data flowing between the business applications of the originator and the recipient become important reporting points that highlight crucial elements of the data's journey between trading partners. Taken sequentially along a path between two partners that use a network interconnect, these milestones are as follows:

Originator System Reporting:

- ❑ Date and time when the originator's business applications routed the transactions to the EDI system
- ❑ Date and time when the originator's application link began preparing the information for translation by the EDI translator
- ❑ Date and time of translation
- ❑ Date and time when the originator's communications software believed it successfully moved the transactions to the originator's EDI network

EDI Network Reporting:

- ❑ Date and time when the originator's network received the transactions from the originator
- ❑ Date and time when the originator's network posted the transactions to the outbound interconnect autodial queue (or in the case of processing or posting failure, the date and time of and reason(s) for the failure)
- ❑ Date and time when the originator's network believed it successfully moved the transactions to the recipient's EDI network
- ❑ Date and time when the recipient's network acknowledged receipt of the interconnected transactions
- ❑ Date and time when the recipient's network (1) failed in its attempt to process and post the transactions to the recipient's mailbox or (2) successfully processed and posted the transactions to the recipient's mailbox
- ❑ Date and time when the recipient retrieved the transactions from the mailbox

Recipient System Reporting:

- ❑ Date and time when the recipient retrieved the transactions from the mailbox
- ❑ Date and time when the recipient translated the transactions
- ❑ Date and time when the recipient interfaced the transactions with their internal business applications
- ❑ Date and time when the recipient acknowledged receipt of the transactions
- ❑ Date and time when the recipient acknowledged (business) acceptance of the transactions

The reporting and conscientious tracking of data movement past these life cycle milestones allows participants to (1) verify timely data movement, (2) identify slowdown points, (3) more finely tune their organizations' leverage of rapid data movement, and (4) detect oddities in the data flow that might be caused by attempts to breech security and gain unauthorized access to data. As discussed previously, when an impostor must not only beat access security to gain access to the data, but also modify control reporting dates and times to avoid detection, the barriers to entry are raised significantly.

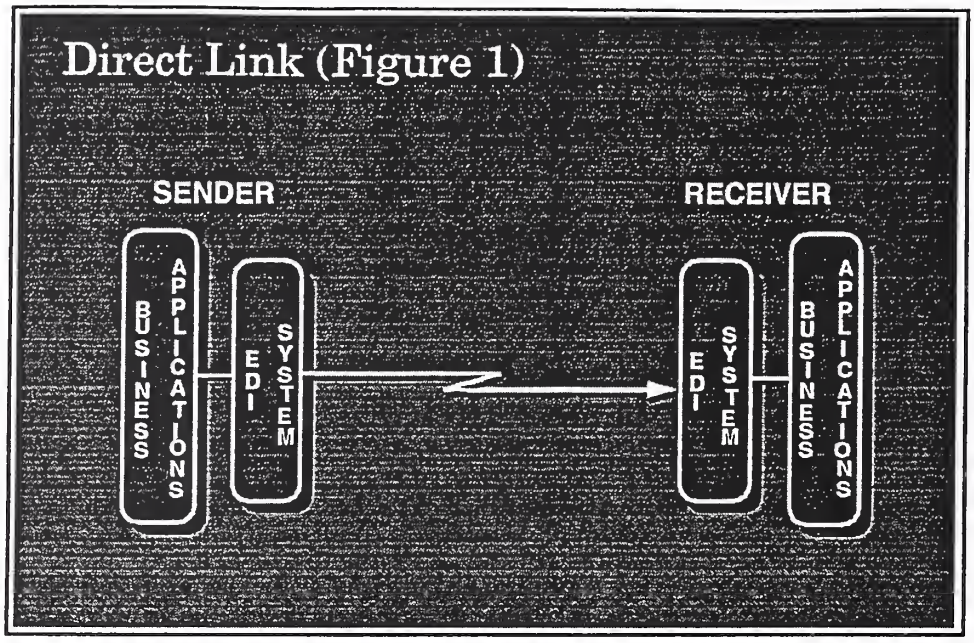
Summary: The Combined Story

The security and controls supplied by the most reputable EDI networks provide measurably improved protection mechanisms over those typically found in direct or point-to-point relationships. The ability to meet all six security and control objectives in a reasonable and, yet, cost-effective manner is decidedly improved. A summary look at the security and control tools available through the use of reputable EDI networks reveals the richness and depth of their impact.

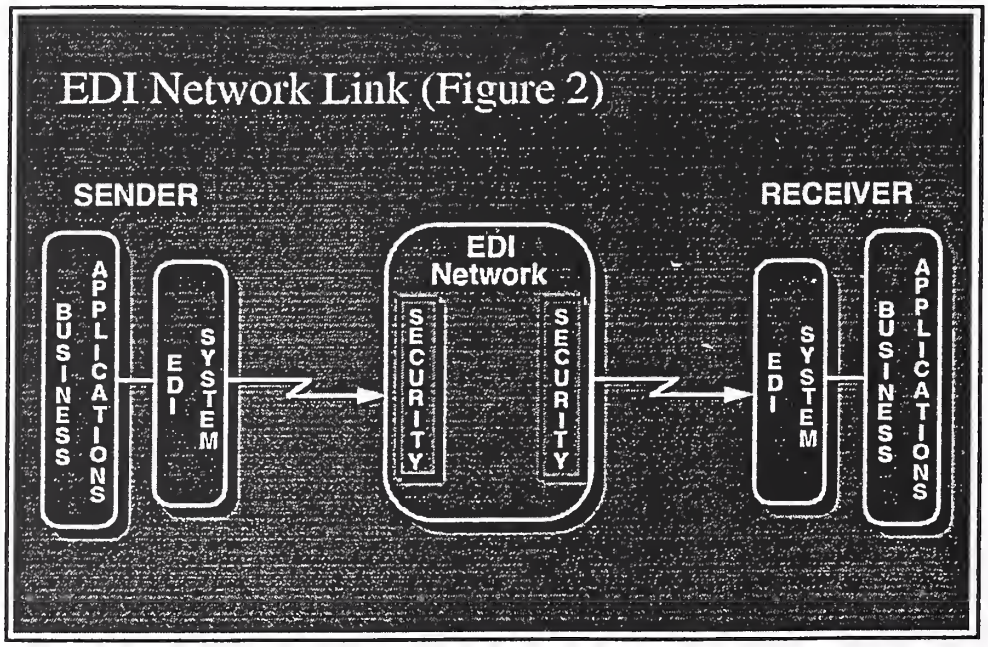
- **Inbound Signon (including signon id and password):** Provides unique identification of the party attempting to access the network services (to send or retrieve data) and specifies a given customer profile to be used by the network to perform further security checks on the party accessing the network and the data being sent and/or received
- **Customer Profile:** Provides specifics on authorized network usage and behavior for a given party. The profile typically includes: signon id and password, special command ids and passwords, communications profile (the protocol, line speed, block size, special configurations such as transparency or compression, etc.), and valid EDI interchange ids that may be used by this party. Customer profiles may additionally indicate valid trading partner pairs for this party, as well as, EDI standards, releases, and transaction set types to be used by this party.
- **Special Network Sessions and Commands (including special command ids and passwords):** Requires accessing party to know the correct commands, sequence of commands, and possibly the correct ids and passwords to be placed in the commands in order to send and/or retrieve data.
- **Confirmation Control Reporting:** Provides the originator with crucial dates and times. The specific dates and times that can be reported are discussed in the previous section (Timely Delivery) under the heading EDI Network Reporting. Common mechanisms for reporting these dates and times are proprietary network control reports (sometimes alternately provided in human-readable, report, format or machine-readable, data, format.) Note: the completion of the X12 Data Status Tracking transaction set (in the near future) will provide a standardized means of providing network control information to end users in a for that can be processed by their translation software.
- **Network Interconnect Receipt Confirmation:** Provides the sending network with pro-active confirmation of receipt and safe storage of data by the receiving EDI network. Common mechanisms for reporting this confirmation are the Interconnect Mailbag Control Structures (including the Interconnect Acknowledgment) or X.400/X.435 acknowledgments.

- **Network Mailbox Posting and Retrieval Notification:** Provides dates and times of mailbox posting and data retrieval to the network(s) of the originator and recipient. This allows the network(s) to report these crucial dates and times to both the originator and the recipient. Common mechanisms for reporting these dates and times are the (soon to be completed) TA3 (for network-to-network reporting) and the DST (for network-to-end party reporting).

Judicious and consistent use of the broad array of security and control tools available in the EDI standards and from EDI networks can provide a high level of reasonable assurance that all six of the security and control objectives are being met in all but the most critical transactions. Only in cases where the value of the data content warrants the extra effort required on the part of impostors to circumvent these barriers is it necessary to resort to the more conclusive (and expensive) security mechanisms such as authentication and encryption.

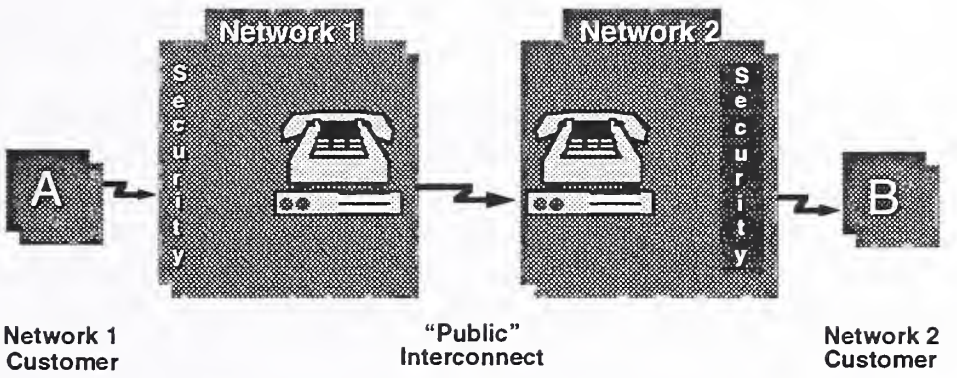


Direct Link (Figure 1)

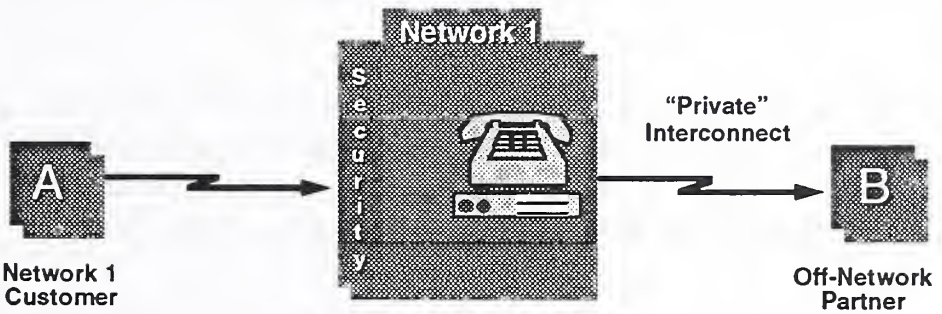


EDI Network Link (Figure 2)

Traditional "Public" Interconnect (Figure 3)



"Private" Interconnect (Figure 4)



Security Requirements and Evidentiary Issues in the Interchange of Electronic Documents: Steps Toward Developing a Security Policy

by Peter N. Weiss[1]

I. Introduction

It is widely expected that the impact of computerization on commerce will be as great as that of the industrial revolution.[2] Electronic messaging techniques, particularly Electronic Data Interchange (EDI), hold great promise to become the preferred methods of communicating administrative and business information. But widespread use of these techniques of electronic commerce will occur only if they have, and are perceived to have, the same or similar level of security as paper-based systems. The concept of security focuses on ensuring the integrity and availability of communications and, to the extent necessary, guaranteeing confidentiality.

Electronic and paper media share many of the same security risks. However, the security protections associated with the traditional use of paper and signatures are so transparent to users and so customary that little thought is given to whether particular transactions require their use. Thus, statutory and regulatory provisions commonly specify that communications be "in writing," "signed," "verified," or "acknowledged." These have become so ubiquitous that most routine paper-based communications, particularly forms, contain a facial requirement for a signature -- even in the absence of any specific legal or administrative directive that an original autograph signature actually be affixed.

In electronic communications environments using techniques such as EDI, however, these security characteristics are no longer "automatic," but must be designed into each particular application. The Computer Security Act of 1987[3] provides a framework for determining what security characteristics are appropriate for particular

---

[1] The views set forth are those of the author and do not necessarily represent those of the Office of Management and Budget.

[2] International Chamber of Commerce, Uniform Rules of Conduct for Interchange of Trade Data by Teletransmission (hereinafter "UNCID Rules"), ICC Publication No. 452 (1988).

[3] Pub. L. No. 100-235, 40 U.S.C. 759 note.

applications. Although the Act only directly addresses Federal computer systems, its principles should be generally accepted. The Act defines sensitive information as including "any information, the loss, misuse, or unauthorized access to or modification of which could adversely affect the national interest or the conduct of Federal programs, or the privacy [of] individuals...." It requires each agency to consider the risk to such sensitive information and to "establish a plan for the security and privacy of each Federal computer system...that is commensurate with the risk and magnitude of the harm resulting from the loss, misuse or unauthorized access to or modification of the information contained in such system."

As the Computer Security Act recognizes, the goal of information security is to manage and minimize risk. The same information which has monetary or other value requiring risk management also may be called upon as evidence to prove the facts associated with particular transactions. A paper purchase order or invoice may become evidence in a contract dispute, and the information in a regulatory filing may be required for enforcement proceedings. In the evidentiary context, the focus is on whether the information is generated in the normal course of business in a manner which maximizes the likelihood that it is reliable and trustworthy.[4] Little consideration has been given, however, to the particular mix of elements which will effectuate the goals of the Act in the EDI context:

> What is needed, then, is a security policy. Various techniques are available to authenticate the source and verify the content of and to control access to electronically transmitted documents. However, there is little jurisprudential guidance as to whether and, if so, under what circumstances these security techniques will provide the requisite assurance of reliability. This lack of guidance concerning security techniques is reflected in the multiplicity of current security and authentication practices within the EDI community.[5]

---

[4] See, U.S. Department of Justice, "Admissibility of Electronically Filed Federal Records as Evidence," (October 1991)(hereinafter "Justice Department guidelines"), reprinted in Information Resources Management Plan of the U.S. Government (December 1991); and "Performance Guideline for the Legal Acceptance of Records Produced by Information Technology Systems," Association for Information and Image Management, AIIM TR31-1992 (hereinafter "performance guideline").

[5] Draft American Bar Association resolution, "Security Techniques in Electronic Transactions," Version 3.0, August 9, 1991. The resulting ABA Resolution did not, however, articulate a substantive security policy, but rather encouraged its development:

II. Steps Toward Developing a Security Policy

The purpose of this paper is to present three preliminary steps toward the development of a security policy for the interchange of electronic documents. Its underlying thesis is that issues of legal admissibility and computer security are intertwined and must be considered together.

First, it briefly reviews basic principles of the law of evidence in order to identify the characteristics of electronic records which maximize the likelihood of their admissibility as evidence. This review suggests that the characteristics associated with the evidentiary value of electronic documents are essentially the same as those associated with maintaining the security of the information. It concludes that the provision of adequate security under the risk-based standard of the Computer Security Act also serves to ensure that the electronic records may be admissible as evidence.

Second, it analyses the security characteristics associated with traditional paper-based communications and compares the functions performed by each with the security services available in electronic data interchange and similar technologies. It demonstrates that although the transition from paper-based communications to electronic techniques poses some unique risks, the essential security requirements are the same.

Finally, it presents a possible security classification scheme for various EDI applications, and suggests presumptively adequate security techniques for each to serve as a starting point for the development of the security plans required by the Computer Security Act and good practice. Each security plan must evaluate the risks associated

---

> The [ABA] supports action by federal and state governments, international organizations, and private entities to: (a) facilitate and promote the orderly development of legal standards to support and encourage the use of information in electronic form, including appropriate legal and professional education; (b) encourage the use of appropriate and properly implemented security techniques, procedures and practices to assure authenticity and integrity of information in electronic form; and (c) recognize that information in electronic form, where appropriate, may be considered to satisfy legal requirements regarding a writing or signature to the same extent as information on paper or in other conventional forms, when appropriate security techniques, practices and procedures have been adopted.

ABA Resolution No. 115 (August 19, 1992).

with the loss, misuse or compromise of the information against the costs associated with the various techniques available to mitigate those risks. Its purpose is to identify those techniques which achieve a reasonable risk/cost balance under the circumstances.

Questions of legal admissibility and computer security are but two sides of the same coin. For example, if a systems manager retained the services of a competent litigator to help design an EDI application in a cost effective manner which would assure a high degree of likelihood that the outputs of the system would be admissible, the system manager would, in the process, have met the requirements of the Computer Security Act. On the other side, had the systems manager retained the services of a security specialist versed in the risk/cost methodology of the Computer Security Act to perform the same task, the outcome should be precisely the reverse -- a high degree of likelihood that the outputs of the application would be admissible as evidence would be assured. Recognition of this essential unity between system integrity and the evidentiary value of system outputs should help to alleviate unfounded, but often expressed, concerns regarding whether electronic documents and their various signature analogues are "legal."[6]

---

[6] Indeed, these concerns should by now have definitively been laid to rest. In general, signature and writing requirements are not legal barriers to electronic commerce:

> **"The concern with electronic signatures...is a red herring.** A variety of techniques for authenticating electronic documents exist that are as good or better than traditional handwritten signatures...There is growing agreement...that authentication and signature concerns can be addressed by existing legal concepts in conjunction with adequate audit and recordkeeping controls."

Perritt, The Electronic Agency and the Traditional Paradigms of Administrative Law, 44 Admin. Law Review 79 (Winter 1992)(emphasis added). See also, ABA Resolution No. 115 (August 19, 1992) footnote 5 supra; "Signature Requirements under EDGAR," Decision of the SEC General Counsel, January 13, 1986 ("Requirements for 'signatures,'... may be satisfied by means other than manual writing on paper...or the use of Personal Identification Numbers (PINs). In fact, the electronic transmission of an individual's name may legally serve as that person's signature, provided it is transmitted with the present intention to authenticate."); and "National Institute of Standards and Technology -- Use of Electronic Data Interchange Technology to Create Valid Obligations," Comp. Gen. Dec. No. B-245714 (December 13, 1991)(Contracts formed using EDI satisfy statutory writing and signature requirements so long as technology used provides same degree of assurance and certainty as traditional "paper and ink" methods of contract formation).

III. Evidentiary Requirements for Electronic Documents

Although the law is sometimes criticized as slow to keep pace with progress,[7] the reality of the information revolution has been recognized by the courts:

> ... [N]o court could fail to notice the extent to which business today depends on computers for a myriad of functions. Perhaps the greatest utility of a computer ... is its ability to store large quantities of information which may be quickly retrieved on a selective basis. Assuming that properly functioning computer equipment is used, once the reliability and trustworthiness of the information put into the computer has been established, the computer printouts should be received as evidence of the transactions covered by the input.[8]

Without going into the details of the application of the "best evidence" and "hearsay" rules,[9] it is sufficient to note that as a general matter computerized records are admissible as evidence provided that they are authenticated and can withstand challenge regarding their genuineness. This authentication requirement is satisfied "by evidence sufficient to support a finding that the matter in question is what its proponent claims."[10] This is done in legal proceedings by "laying a foundation" that will qualify the evidence as being what is purported to be (e.g. a record prepared in the ordinary course of business).

Electronically filed Federal records are almost invariably offered as business records prepared in the ordinary course of business. During the process of laying the foundation, the proponent of the evidence seeks to demonstrate the authenticity and reliability of the information, and the opponent tries to challenge those assertions:

> ... [T]he foundation for admission of (computerized records)consists of showing the input procedures used, the tests for accuracy and reliability and the fact that an established business relies on the computerized records in the ordinary course of carrying on its activities. The (opposing) party then has the opportunity to cross-examine concerning company practices with respect to the

---

[7] See Weiss, "Law and Technology: Can They Keep Abreast?," 8:4 Government Information Quarterly 377 (Nov. 1991).

[8] Harris v. Smith, 372 F.2d 806 (8th Cir. 1967)(emphasis added).

[9] These are discussed in detail in the Justice Department guidelines, supra note 4.

[10] Federal Rule of Evidence 901(a).

> input and as to the accuracy of the computer as a memory bank and retriever of information....[T]he court (must) "be satisfied with all reasonable certainty that both the machine and those who supply its information have performed their functions with utmost accuracy." ... [T]he trustworthiness of the particular records should be ascertained before they are admitted and... the burden of presenting an adequate foundation for receiving the evidence should be on the parties seeking to introduce it rather than upon the party opposing its introduction.[11]

Federal records management regulations also incorporate these evidentiary principles and provide similar guidance to Federal agencies in carrying out their responsibilities.[12]

In sum, the law of evidence does not rest on inflexible paper-based rules which should pose a barrier to the use of electronic commercial practices. Rather, it is concerned with the underlying integrity of the information on which a judge or jury can reasonably rely in reaching a just conclusion to a particular controversy. Modern rules of evidence and court decisions appear to have come to terms with the realities of business and professional practice -- the ever-growing dependence on information technology systems for records production and maintenance.[13]

The essential questions posed by the law of evidence in this context can be summarized as follows:

> Electronic messages present four distinct evidentiary problems:
>
> 1. Proving that an electronic communication actually came from the party that it purports to come from;
>
> 2. Proving the content of the transaction, namely, the communications that actually occurred between the parties during the contract formation process;
>
> 3. Reducing the possibility of deliberate alteration of the contents of the electronic record of the transactions;

---

[11] United States v. Russo, 480 F.2d 1228 (6th Cir. 1973).

[12] "Judicial Use of Electronic Records," at Par.11, Federal Information Management Regulation (FIRMR) Bulletin B-1, "Electronic Records Management," and at 36 C.F.R. Part 1234.24.

[13] "Performance guideline" supra note 4 at Section 2.4.

4. Reducing the possibility of inadvertent alteration of the contents of the electronic record of the transactions.[14]

The key evidentiary issue is the weight that a court will give to electronic records. "This will primarily be a question of agreeing to, and implementing, adequate security procedures."[15] These concerns with the identification of the originator, the integrity of the content of the communication, and with reducing the likelihood of alteration, which are at the heart of the law of evidence, are precisely the concerns which must be addressed in the context of EDI security. Thus, any combination of security controls which provides assurance that these characteristics have not been compromised will also provide a high degree of confidence that the contents of the communications will be admissible as evidence. The following sections of this paper will examine the security characteristics associated with both paper and electronic media, and will suggest a security classification scheme which may be of assistance in determining the appropriate mix of security techniques for particular EDI applications.

IV. Security Characteristics of Paper-Based Communications

Traditional paper-based communications accompanied by handwritten signatures provide three essential security characteristics: **message integrity, originator authentication**, and **non-repudiation**. Depending on the nature of the communication, an additional security characteristic, **confidentiality**, may be desired. The efficacy of the various techniques used to ensure the desired level of security in turn depends on the adequacy of the **administrative controls** associated with their use.

Message integrity is the assurance that the content of a communication is complete and has not been changed prior to receipt. This is accomplished by a number of features, the primary ones being those associated with the use of writing itself: inks which make erasure and alteration easily perceptible, salutations and closings which constrain the length of the message, and even the size of the paper which may limit the addition of text. For applications requiring additional security, techniques such as the use of engraved backgrounds, chemically treated papers or lamination in plastic are used to make alteration particularly difficult.

---

[14] Baum & Perritt, Electronic Contracting, Publishing and EDI Law, Section 6.23, "Evidentiary Issues" (Wiley 1991).

[15] Edwards, Ed., Information Technology and the Law (Second Edition)(Macmillan, 1990) at p. 241.

Originator authentication provides assurance that the communication originated with the named source. This is most commonly provided by the handwritten signature or, historically, by the seal of the author. The authentication purpose of the signature or the seal has two conceptual parts: First, they add a degree of formality, increasing the likelihood of actual assent to the terms contained in the document. Second, they serve to identify the document with the originator, because signatures and seals tend to be unique. In most commercial transactions today, these functions are served primarily by the use of letterhead or pre-printed forms, and in formal documents of a routine nature such as checks and negotiable instruments printed signatures or "autopens" are often used to fulfill legal and customary requirements for signatures. Higher levels of originator authentication can be provided by the use of watermarked or other special paper such as those generally used for negotiable instruments.

Non-repudiation is a stronger form of authentication which relates to the ability of a disinterested third party reasonably to conclude that the identified originator intended to be bound by the substance of the communication. This function is most commonly performed by the original autograph signature affixed to a document having facially adequate message integrity. During the early development of contract law, primary reliance was placed on the individual's seal as indicating intent to be bound. Only in the 20th Century did the signature gain its present prominence, and the special status accorded contracts under seal only disappeared as states enacted the Uniform Commercial Code. Enhanced forms of non-repudiation have generally involved the use of witnesses. Even after the use of written records of business and other transactions became common in the late Middle Ages, the most important enhanced form of non-repudiation remained witnesses. This formal reliance on witnesses is carried over today in the attestation of wills and the use of notaries public.

Confidentiality is the ability to limit access to the information contained in a communication. This has generally been accomplished with some combination of security markings, envelopes, seals, trusted messengers, and by the use of codes and ciphers.

Central to the efficacy of message security is the use of adequate **administrative controls**. As paper-based communications took on forms more diverse than the handwritten document with affixed signature, communicating entities had to establish internal procedures to assure the efficacy of the various security techniques they wished to utilize. These ranged from limiting access to the official seal, and later to

the letterhead and autopen, to ensuring the trustworthiness of message carriers and witnesses.[16]

As explained below, these same security characteristics and procedures are associated with electronic communications, particularly EDI. The primary difference is that the ubiquity of these techniques in paper-based systems and their transparency to users results in their being given little attention. It is generally only when cost becomes a relevant factor -- e.g. the costs associated with special papers, autopens, or bonded couriers -- that attention is given to the risk/cost equation. In EDI systems, however, no intrinsic security "baseline" analogous to the forensic characteristics provided by paper, ink and signatures exists. Rather, each technique as applied to any particular application carries a price.

V. Security Characteristics of Electronic Data Interchange

> "...each time a new system or tool is produced, our more or less conscious attachment to tradition leads us to expect guarantees which were previously not only never fulfilled but were not even asked for."[17]

The use of electronic commerce techniques does not necessarily increase transactional risk beyond that experienced in a paper-based environment, but in some ways can actually reduce the likelihood of legal disputes ever arising. This is in spite of the fact that, unlike paper-based communications, electronic communications can be changed without a trace. For example, some security techniques are built into relevant communications protocols (e.g. X.25 and X.400) and the EDI standards themselves contain headers, password fields and control information. These characteristics, coupled with the speed of communication afforded by EDI and the decreased likelihood of transcription errors, may well lessen the frequency of disputes caused by the transmission of erroneous information.

A common-sense corollary to the risk-based standard set forth in the Computer Security Act is that, except where the use of computers increases risk, the use of computers should not create new requirements for the conduct of business beyond those that exist in a paper environment, unless the additional security obtained from

---

[16] OMB Circular No. A-123, Internal Control Systems, sets forth the present requirements for administrative controls in Federal agencies.

[17] A.A. Martino, quoted in Edwards, supra, p. 241.

those measures is worth the additional cost.[18] Looked at from the standpoint of potential threats, "[s]uch controls should make the cost of obtaining data greater than the potential value of obtaining or modifying the data."[19]

Security characteristics of paper-based media -- hand or typewritten signatures and letterhead -- are relatively easy to defraud, yet we use them unless the particular transaction is of such value that the cost of additional precautions seems justified. Certain electronic techniques can provide security beyond that available in a paper environment, and should be used when they will cost-effectively control new or previously uncontrollable risks. The point is that security is not an absolute, but must be tailored to the particular circumstances in order to be "commercially reasonable."[20]

As in a paper-based system, the use of appropriate administrative controls is essential to assuring security in EDI applications. These include organizational arrangements such as separation of duties, physical security, and techniques for message authorization. Adequate administrative controls are central to the ability effectively to make use of the various techniques available to ensure the requisite level of security in the particular EDI application. In an electronic commerce environment, administrative controls also must be agreed upon and followed by trading partners. The International Chamber of Commerce's UNCID rules set forth a code of conduct under which trading partners may agree on such factors as appropriate identifiers, acknowledgments of transactions, confirmation of contents, protection of sensitive data, and data storage and transaction logging.[21]

The following is a description of the various computer security techniques applicable to EDI, followed by an indication of which of the four security characteristics, discussed above, they tend to satisfy. They are listed in a generally ascending order

---

[18] McConnell, "Electronic Data Interchange in the U.S. Government: An Active Ingredient of Electronic Commerce," 1991:1 EDI Forum 17, reprinted from A Five Year Plan for Meeting the Automatic Data Processing and Telecommunications Needs of the Federal Government (November, 1990).

[19] ISO 7498, Addendum 2.

[20] See Uniform Commercial Code, Article 4A.

[21] Note 2 supra.

of security strength.[22] However, the strength of each technique depends on how it is integrated into the system and the accompanying administrative controls.

o **Access controls**. The use of logon techniques including passwords, key cards or other tokens, remote job entry protocols or other unique identifiers such as fingerprint configuration or other biometric characteristics, which identify users and restrict access to an EDI application. **Originator authentication, non-repudiation, confidentiality.**

o **Imbedded references**. The use of agreed reference numbers or passwords, either generic to the parties or specific to particular transactions, within a message. **Originator authentication, non-repudiation.**

o **Functional acknowledgment**. A requirement for a confirmation message to be returned each time a message is received, but which does not repeat back the contents of the message. Analogous to a postal return receipt. **Originator authentication, non-repudiation.**

o **Message repetition acknowledgment**. A requirement for a confirmation message to include the full contents or critical elements of the message sent. **Message integrity, originator authentication, non-repudiation**.

o **Internal message verification**. Recalculation and verification of real totals and/or hash totals to protect against altered values of essential fields of a message. Hash totals are summations for checking purposes of similar fields, such as those containing part numbers, which would otherwise not be summed. **Message integrity**.

o **Trusted Third-Party**. The use of a third-party service provider or value added network (VAN) to provide message status reports, message filing and audit services, and other security services. **Originator authentication, message integrity (depending on service), non-repudiation, confidentiality**.

o **Cryptographic message authentication**. Techniques which utilize message authentication codes (MAC) and "digital signatures" calculated from all bits in

---

[22] Additional discussion of each is contained in the NIST Computer Systems Laboratory Bulletin "Security Issues in the Use of Electronic Data Interchange" (June 1991), and the references cited therein.

the message using a secret encryption key. May be verified, if desired, by a recipient having possession of a decryption key. See, e.g. FIPS PUB 113. **Originator authentication, message integrity, non-repudiation.**

o **Data encryption**. Encrypts all bits of a message. Keys used for confidentiality must be different than those used for cryptographic message integrity, and both parties must have key access. See, FIPS PUB 46-1. **Confidentiality, non-repudiation.**[23]

These security techniques and their functions are summarized in Table 1. Their relative strengths are not indicated in the Table. They comprise a menu of techniques which, alone or in combination, can provide various levels of security.[24] Each, however, has its cost in terms of administrative effort as well as the hardware and software needed for their implementation. The mapping of security techniques to function is for general guidance and is based on expected usage. For instance, access controls and message authentication techniques can provide non-repudiation if the techniques are strong and the application supports it. Conversely, cryptography without administrative controls may not provide non-repudiation.

---

[23] The practical and cost implications of the use of public key versus private key cryptosystems for message authentication and data encryption are beyond the scope of this paper.

[24] This listing is generally consistent with that contained in the ABA's "Model Payments Agreement and Commentary," 32 Jurimetrics Journal No. 3 (1992), which lists verification techniques of generally ascending strength:

"(a) Sequence number consistency;
(b) Comparison of control totals with Remittance Information associated with a payment;
(c) Use and confirmation of a valid password/user ID combination;
(d) Communication call back procedures, and use of private or leased communication lines;
(e) A syntactical check on the Transaction Set as received, together with the subsequent communication of a Functional Acknowledgment to the Transaction Set's originator;
(f) Consistency checking of the payment amount with prior transactions or customer profiles;
(g) Smart cards and 'tokens;'
(h) Message Authentication Codes; and
(i) Digital signatures."

VI. Presumptive Security Levels for Various Electronic Data Interchange Applications

The security assessment associated with each EDI application should include an examination of the substantive nature of each transaction type and an analysis of the risks and threats associated with each. Applications range from those which are not sensitive (e.g. reports of order status or questionnaires involving information without privacy or business confidentiality implications), through those with low to medium levels of sensitivity (e.g. procurement transactions and regulatory reporting), to those with high sensitivity (e.g. electronic funds transfer). The desired mix of security techniques will differ for each.

It is also important to recognize that it is the substance of the transaction rather than its form which is critical to the security analysis as well as to the issues associated with legal admissibility.[25] For example, an EDI purchase order which is of relatively low dollar value and which is part of a routine course of dealings between trading partners would likely have a low level of risk from tampering or other threats. Likewise, it would require a relatively straightforward foundation for admissibility as evidence in the event of a dispute. On the other hand, an identically formatted EDI purchase order which is of a high dollar value and exchanged between parties who have never done business before would likely have a higher level of risk from tampering or repudiation. It would require a more extensive foundation for admissibility in evidence.

Since this analysis focuses on the security of the data interchange process, it does not examine a related issue relevant to admissibility: the security of the storage of messages after receipt. One of the keys to laying a proper evidentiary foundation is the ability to demonstrate that an organization's recordkeeping practices are such that

---

[25] Contrary to some popular belief, use of encrypted message authentication techniques is not necessary to satisfy legal "signature" requirements. See fn. 6 supra. For example, the Comptroller General's opinion suggests that an electronic signature must be "bound" to the data, and that only encryption can fulfill that requirement. However, signatures in paper media are not "bound" to the data content in the same manner as encryption, but are merely "affixed" to the paper, sometimes even before the data is written. Furthermore, the legal literature is devoid of any reference to a "binding" requirement. Rather, encryption is but one of a number of techniques that can satisfy signature requirements.

their outputs can be deemed credible reflections of their inputs.[26] Thus, the evidentiary showing regarding records security may also vary based on media and storage techniques. For example, it is likely that electronic records stored on write-once-read-many (WORM) optical media may be considered to have a higher degree of security, and hence be more readily admissible, than records stored on magnetic media.[27] The security characteristics of an organization's data storage methods must, of course, be considered as part of the overall security analysis:

> Good electronic record systems design ensures that archives and records retention needs are designed into the system. While such design features may be difficult to incorporate in PC-based systems, the communications link...is an obvious and fail-safe point of capture for maintaining a comprehensive record....Technical means could ensure that nothing gets into the system without being entered in a docket and having an archival copy made. The integrity would be greater than that achievable with human- and paper-based systems.[28]

The following schema is intended to aid in security analyses of EDI applications. It sets forth four general categories each with increasing levels of security requirements, suggests a mix of security techniques presumptively appropriate for each level, and provides examples of applications which generally would be considered to be in the particular security category.

> o **Non-Sensitive**. Applications which do not involve the obligation of Federal funds and which do not have regulatory or privacy implications. Examples include order status information, material inspection and receiving reports, and some questionnaires. For these applications, reasonable access controls should be adequate with other techniques optional.
>
> o **Sensitive (Low)**. Applications which have no significant incentive for tampering by third-parties. These include most small purchase transactions, orders, invoices, bills of lading, and most regulatory reporting applications. Originator authentication and non-repudiation can generally be satisfied by functional acknowledgments, and the risk of tampering and privacy concerns, if any, can be minimized through access controls. Additional authentication and

---

[26] See, Justice Department guidelines, supra note 5.

[27] Id. at note 2.

[28] Perritt, "Electronic agency" supra note 6.

non-repudiation techniques such as message repetition, internal message verification, and imbedded references are optional.

o **Sensitive (Medium)**. Applications which present significant incentives for tampering and/or for which a reasonable level of confidentiality should be maintained. These include responses to Invitations for Bids and Requests for Proposals as well as applications for valuable benefits or substantial payments. Either of two strategies may be used. Cryptographic data authentication as described in FIPS PUB 113 or similar techniques provide strong protection against tampering. The use of message repetition acknowledgment, or other message verification techniques, in conjunction with a trusted third-party service provider may be adequate provided that the service provider has strong system access controls and adequate recordkeeping and audit mechanisms.

o **Sensitive (High)**. Applications where message confidentiality is of particular concern, or where there is a particularly great risk from lack of message integrity, and access related controls are deemed inadequate. These include the protection of particularly sensitive though unclassified information such as electronic funds transfer transactions. Generally, encryption techniques are recommended, either full text encryption for confidentiality or cryptographic message authentication.

These sensitivity levels and their presumptive security techniques, along with examples of each, are summarized in Table 2. It should be noted that particular transactions may have varying levels of sensitivity for differing parameters. For example, while encryption may be considered appropriate for an electronic funds transfer, the remittance advice information related to the transfer may have a low degree of sensitivity for originator authentication and message integrity. Depending on the nature of the transaction, there may or may not be confidentiality concerns. Thus, the analysis may at times be multi-dimensional.

Dealing with confidentiality concerns is particularly challenging. On the one hand, only cryptographic techniques can ensure a high degree of confidentiality. However, in paper-based systems, the business community has accepted that the confidentiality provided by the postal system is adequate and that the risk of their information being improperly divulged is acceptably low. Therefore, it may be that the risk of such disclosure on electronic networks, absent the use of encryption, is also acceptably low. This depends on an analysis of the strength of the access controls related to the system and the type of transaction.

This may be the case since the private sector routinely transmits confidential business information unencrypted. While it is certainly possible for data to be intercepted while it is on a vendor's network, it is more likely to be improperly accessed while it is still in the hands of the company. When data spies use telecommunications networks, it is usually to gain access to a company's computers.

If parties to particular transactions think that the risk of disclosure from unsecured telecommunications links is too high, then additional levels of security can be added. While the installation costs of a data encryption capability may be low, the maintenance costs (especially at the administrative level) may be an impediment to the use of this technology, at least in the near term. Moreover, data encryption is not now in wide use for commercial transactions other than funds transfer. Careful attention must be paid to the risk/cost tradeoffs in these situations.

VI. Conclusion

The thesis of this paper is that evidentiary issues and security requirements are two sides of the same coin. And in the realm of security "one size" does not fit all, just as in the law of evidence the foundational showing will vary with the particular circumstances.

A simple hypothetical problem should elucidate the point. Party A sends Party B an electronic purchase order in standard EDI format. Parties Y and Z do the same. In both cases disputes arise necessitating the use of the two purchase orders as evidence. Here, however, the similarities end. Parties A and B, it turns out, are established trading partners engaged in a regular course of business involving the routine exchange of electronic purchase orders. The transaction at issue involved a standard commercial product and did not carry an extraordinary dollar value. Parties Y and Z, however, are strangers who -- although they possess and utilize EDI capabilities -- have never done business together before. Furthermore, the transaction was of a high dollar value and was for the purchase of a custom manufactured item.

Although the two EDI purchase orders were essentially identical, from an evidentiary standpoint the two transactions were totally different. The burden party A must carry in order to have its purchase order admitted into evidence is relatively light. The use of basic security techniques -- password access control, generally reliable audit capability, probably the use of a VAN -- should suffice to have the evidence admitted. Party Y, however, must bear a heavy evidentiary burden. The controls used by party A would probably not suffice. Strong originator authentication, message integrity, and non-repudiation -- probably encryption techniques -- should have been used.

Likewise, from the standpoint of the Computer Security Act's risk-based standard, the two transactions bear little resemblance. For parties A and B, use of sophisticated and potentially costly security techniques as a supplement to routine control and audit practices would have been unnecessary to satisfy the Act. For parties Y and Z, they would probably have been essential.

In sum, the development of security plans as required by the Computer Security Act and good practice involves a common sense approach to risk assessment. Analyzing the security requirements of particular applications can be aided by considering the security characteristics which the application should possess as well as the sensitivity level for each. As enhanced security techniques become more cost effective and increasingly ubiquitous, the task will become easier. However, careful assessment of the risk/cost tradeoffs must be made as part of this process. Attention to these factors should satisfy applicable legal requirements.

Table 1

Document Interchange Security Techniques and Characteristics

SECURITY TECHNIQUES	CHARACTERISTICS			
	Originator Authentication	Confidentiality	Non-repudiation	Message Integrity
Access Controls*	X	X	X	
Imbedded References	X		X	
Functional Acknowledgment	X		X	
Message Repetition Acknowledgment	X		X	X
Internal Message Verification				X
Trusted Third-Party	X	X	X	X
Cryptographic Data Authentication	X		X	X
Data Encryption		X	X	

* Access controls include a variety of techniques providing a wide range of security strength.

Sensitivity Level and Presumptive Security Techniques	Application Examples	
	Procurement	**Non-Procurement**
Sensitive (High)		
Text Encryption per FIPS PUB 46-1	Protection of Logistics Unclassified Sensitive (PLUS)	Regulatory and other reporting with particular confidentiality concerns
EFT encryption	Electronic Funds Transfer	Electronic funds transfer
Sensitive (Medium)		
Cryptographic message authentication or Message repetition acknowledgment through trusted third-party	Invitations for Bids Requests for Proposals	Regulatory and other reporting with significant incentive for third-party tampering
Sensitive (Low)		
Access controls, Functional acknowledgment Optional: Message repetition, Internal message verification, Imbedded references	Requests for Quotations Orders under existing contracts Invoices Government Bills of Lading	Regulatory and other reporting: tax filings customs filings environmental reports Personnel actions
Non-Sensitive		
Access controls	Bidders mailing list information Status of orders Reports on orders received Material inspection and receiving reports	Questionnaires without confidential or proprietary information