2001

The Emperor's New Clothes: The Shocking Truth About Digital Signatures and Internet Commerce

Jane K. Winn
University of Washington School of Law

Recommended Citation

Jane K. Winn, *The Emperor's New Clothes: The Shocking Truth About Digital Signatures and Internet Commerce*, 37 Idaho L. Rev. 353 (2001), https://digitalcommons.law.uw.edu/faculty-articles/167

THE EMPEROR'S NEW CLOTHES: THE SHOCKING TRUTH ABOUT DIGITAL SIGNATURES AND INTERNET COMMERCE

JANE K. WINN*

TABLE OF CONTENTS

I. INTRODUCTION: THE HYPE SURROUNDING DIGITAL SIGNATURES 354
II. THE ORIGINAL CONSENSUS: DIGITAL SIGNATURE AS SIGNATURE 362
A. Does the Metaphor of "Signature" Make Sense for Asymmetric Cryptography and Public Key Infrastructures? 364
B. Why Do Signatures Matter in Traditional Contracting Practices? 367
C. What Does "Non-Repudiation" Mean? 370
III. COMMERCIAL APPLICATIONS OF DIGITAL SIGNATURE TECHNOLOGY 375
IV. LAW REFORM AND AUTHENTICATION IN ELECTRONIC COMMERCE 379
V. CONCLUSION 382
VI. APPENDIX: ASYMMETRIC CRYPTOGRAPHY, DIGITAL SIGNATURES AND PUBLIC KEY INFRASTRUCTURE 384

> So off marched the emperor in the procession under the beautiful canopy, and everybody in the street and at the windows cried: "Aren't the emperor's new clothes wonderful!

[*] Professor, Dedman Law School, Southern Methodist University, Dallas TX; coauthor, The Law of Electronic Commerce, 4th ed (2001). I would like to thank Tom Albertson, Dwight Arthur, Thomas Blackwell, Nicholas Bohm, William Boyd, Todd Boyle, Stefan Brands, Roger Clarke, Paolo Da Ros, Don Davis, Jered Floyd, Patricia Fry, Maureen Garde, John Gregory, Ian Grigg, David Goodenough, Harald Hanche-Olsen, Robert Hettinga, Rick Hornbeck, Ted Janger, Dale Johnson, Lyn Kennedy, Jay Kesan, Ben Laurie, Neal McBurnett, John Messing, John Muller, Christina Ramberg, Scott Renfro, Greg Rose, Jeffrey Rothstein, Joseph Sommer, Gary Stock, Simone van der Hof, Richard Warner, Benjamin Wright, and Jonathan Zittran for providing me with valuable feedback on earlier drafts of this article. Any remaining errors are my responsibility alone.

What a lovely train he has to his robe! What a splendid fit!" Nobody would let on that he couldn't see anything, because then he would have been unfit for his job or very stupid. Never had the emperor's clothes been such a success.

"But he hasn't got anything on!" cried a little child.

"Dear me! Listen to what the pretty innocent says!" cried its father. And it was whispered from man to man what the child had said.

'"He hasn't got anything on," says a little child. "He hasn't got anything on!"'

"Why, but he hasn't got anything on!" they all shouted at last. And the emperor winced, for he felt they were right. But he thought to himself: "I must go through with the procession now." And he drew himself up more proudly than ever, while the chamberlains walked behind him, bearing the train that wasn't there.

The Emperor's New Clothes, Hans Christian Andersen, translated by Reginald Spink (1960).

I. INTRODUCTION: THE HYPE SURROUNDING DIGITAL SIGNATURES

It has been an article of faith for several years now among many observers that digital signatures[1] will be the "next big thing" for

[1] This article follows what is now a widely followed convention in electronic commerce circles by referring to a specific application of a specific technology as a "digital signature" and using the term "electronic signature" to refer to electronic authentication technologies that serve the same purpose as manual signatures. In this context, a digital signature refers to:

> [a] transformation of a message using an asymmetric cryptosystem and a hash function such that a person having the initial message and the signer's public key can accurately determine (1) whether the transformation was created using the private key that corresponds to the signer's public key, and (2) whether the initial message has been altered since the transformation was made.

Information Security Committee, A.B.A. Sec. Sci. & Tech., DIGITAL SIGNATURE GUIDELINES: LEGAL INFRASTRUCTURE FOR CERTIFICATION AUTHORITIES AND SECURE ELECTRONIC COMMERCE § 1.11 (1996) [hereinafter DIGITAL SIGNATURE GUIDELINES]. By contrast, an electronic signature may refer to a name in the "From" header in an electronic mail message, a digitized handwritten signature such as are used by some retail

Internet commerce.[2] Digital signatures, authenticated with reference to certificates administered within a "public key infrastructure" may hold tremendous promise as a solution to the problem of establishing the identity of parties doing business in cyberspace. That unrealized potential is consistently mistaken for actual use in the marketplace, however, leading to countless wildly inaccurate journalistic accounts of digital signatures as the "most popular" or "most important" system for Internet contract formation.[3]

Yet in early 2001, the number of Internet contracts that were being formed in reliance on digital signature certificates still appears to be trivially small in number, if not actually zero.[4] Furthermore,

electronic point of sale payment systems, or a typed electronic version of a paper-based holographic signature such as "/s/Jane Winn." *Id.*

The commercial value of a digital signature is usually presumed to be a function of the certificate issued attesting to the identity of the owner of the digital signature. In the Digital Signature Guidelines, a function served by a certificate is explained in the following terms:

> To associate a key pair with a prospective signer, a certification authority issues a certificate, an electronic record which lists a public key as the subject of the certificate, and confirms that the prospective signer identified in the certificate holds the corresponding private key. The prospective signer is termed the subscriber. A certificate's principal function is to bind a key pair with a particular subscriber.

DIGITAL SIGNATURE GUIDELINES at 13.

[2] *See* Carl Ellison and Bruce Schneier, *Ten Risks of PKI: What You're not Being Told About Public Key Infrastructure*, 16 COMPUTER SECURITY J. 1, 1-7 (2000), *available at* http://www.counterpane.com/pki-risks.html; Don Davis, *Compliance Defects in Public-Key Cryptography*, Proc. 6th Usenix Security Symp, 171-178 (San Jose, CA, 1996), *available at* http://world.std.com/~dtd/compliance/compliance.ps. *But see* Ben Laurie, *Seven and a Half Non-risks of PKI*, *available at* http://www.apache-ssl.org/7.5things.txt

[3] *See, e.g.*, Sheryl Canter, *Electronic Signatures – Now it's legal to sign documents electronically, but should you?*, PC MAG., Jan. 2, 2001, at 102, *available at* Lexis News ("The most common technology used for electronic signatures is the digital signature."); Leslie Brooks Suzukamo, *E-Signatures Gain Force of Law, But Users Face a Learning Curve*, ST. PAUL PIONEER PRESS, Oct. 1, 2000, *available at* Lexis News ("In its most common form, a digital signature is, quite simply, extremely long strings of numbers and letters put together by a mathematical formula."); James K. Watson, Jr. and Carol Choksy, *Digital Signatures Seal Web Deals*, INFORMATIONWEEK, Sept. 18, 2000, *available at* Lexis News ("Digital signatures can be any form of electronic seal agreed to by the two parties. The most common approach relies on digital certificates and encryption."); Thomas E. Crocker, *Resolve State Conflicts with Federal Electronic Authentication Law*, LEGAL TIMES, Mar. 1, 1999, at S43 *available at* Lexis News ("The most widely accepted form of electronic authentication currently is based on cryptographic measures, such as digital signatures, which involve mathematical formulas.")

[4] The figure of zero Internet contracts formed in reliance on digital signatures may be accurate if pilot projects are excluded. *See, e.g.*, Tony Heffernan, *Digital Signatures Still 3 to 5 Years Away*, THE AM. BANKER, Jan. 8, 2001 at 2A, *available at* Lexis News; Jamie Lewis, *PKI Won't Hit the Mainstream Until Vendors Reduce Complexity*,

there is no indication that the situation will suddenly change in the near future. After many years of enduring mind-numbingly dull explanations of asymmetric cryptography, hash functions, public key infrastructures and stories of Bob and Alice who want to communicate with the assistance of Carol certificate authority,[5] perhaps the time has come to admit that the market reality has not matched the hype. This might also be a good time to analyze how the enthusiasm for this technology could have reached such feverish heights in the absence of any significant use in the marketplace, and how that enthusiasm can persist today in the face of fairly compelling evidence that the hype will never be realized.

This Article critiques a specific set of assumptions about specific application of digital signature technology: that contracts will be formed over the Internet among parties with no prior relationships through reliance on digital signature certificates issued by trusted third parties to establish the identity of the parties. This application for digital signature technology was once seen as both its most ambitious and most promising application because, for parties with no prior knowledge of each other, there is not yet a reliable system of online identities in Internet commerce. Parties with an ongoing commercial relationship can absorb the cost of offline communications such as faxes, telephone calls or face-to-face meetings to negotiate and execute an agreement governing the setting up of a reliable system for online authentication of parties to wholly electronic transactions.[6] Parties that want to rely exclusively on online communications to create the framework for contracting as well as to enter into contracts, however, face a problem of infinite regress: how can the online communications that set up the system for confirming online identities it-

INTERNETWEEK, Jan. 8, 2001 at 25, *available at* Lexis News; Kelly Jackson Higgins, *Public Key Infrastructures - Few and Far Between*, INTERNETWEEK ONLINE (Nov. 2, 2000), *at* http://www.internetweek.com/lead/lead110200.htm; Tara C. Hogan, *Now That the Floodgates Have Been Opened, Why Haven't Banks Rushed Into the Certification Authority Business?*, 4 N.C. BANKING INST. 417 (2000); *Digital Certificates: A Solution in Search of a Problem*, March 8, 2001 (pointing out "As an example, consider the experience of VeriSign Inc., a leading manufacturer of the 128-bit devices. Since 1998 it's installed more than 25,000 server certificates to financial institutions, but only about 345 client certificates."), *available at* http://www.thebankingchannel.com/technology/story.jsp?story=TBCQFNQUI JC. In a February 14, 2001 email to the author, Ian Grigg contested the zero figure, and pointed to http://webfunds.org/ricardo/contracts/digigold/ as an example. This Internet contracting system is based on OpenPGP's web of trust, not a hierarchical PKI.

[5] *See generally* Jane K. Winn, *Couriers Without Luggage: Negotiable Instruments and Digital Signatures*, 49 S.C. L. REV. 739, 763 n.150 (1998).

[6] *See generally* ABA Electronic Messaging Services Task Force, *The Commercial Use of Electronic Data Interchange: A Report and Model Trading Partner Agreement*, 45 BUS. LAW. 1645 (1990).

self be authenticated with nothing more to rely on than online communications? Many supporters of digital signatures believed legislation was essential to cut through this Gordian Knot. Legislation could authorize parties unable to use a prior relationship or offline communications to confirm the validity of online identities to rely on digital signature certificates instead. Much legislation regulating the use of digital signatures is based on an unstated premise: liabilities must be imposed by law because private agreements will not be adequate to the task of regulating this technology.

What is now becoming apparent is that a more important commercial for digital signatures than "open" Internet commerce among strangers may be "closed" Internet commerce systems among parties already in contractual privity with each other or to a system administrator. Internet commerce systems are being developed that require parties to subscribe to a common, binding set of system rules, or lower transaction costs among parties with pre-existing relationships. If this is the "killer application" for digital signature technology, then laws drafted to facilitate transactions among parties without a common set of system rules or prior relationship may actually interfere with the ability of interested parties to build and operate effective online contracting systems.[7] Because the Uniform Electronic Transactions Act (UETA) is silent with regard to what technologies parties may use in electronic transactions and how they use those technologies, it is unlikely to impose any unnecessary transaction costs on parties to electronic contracts.

In addition to "closed" Internet commerce systems, another major application for digital signature technology may exist within network security infrastructures that are transparent to transacting parties as they operate. Secure Sockets Layer (SSL) communication security is an example of such an application.[8] Such an application diverges from the specific set of assumptions about the "signature" function of digital signatures this Article critiques because there is no conscious invocation of the technology to "sign" anything by the end user. Similarly, this Article does not address the use of digital signature certificates as a substitute for sign-on systems that today rely on user IDs and pass-

[7] For example, it is unclear the scope of the parties' ability to opt out of the liability allocation provisions in the UNICTRAL Draft Model Law on Electronic Signatures. *See*, A/CN.9/WG.IV/WP.88 - Draft Guide to Enactment of the UNCITRAL Model Law on Electronic Signatures *available at* http://www.uncitral.org/english/workinggroups/wg_ec/wp-88e.pdf; Stewart Baker, et al., UNCITRAL Working Group Approves Model Law on Electronic Signatures (November 2000 draft memo).

[8] See *infra* Part III for a discussion of Secure Sockets Layer technology.

words, or government programs for distributing digital signature certificates to control access to government records or to permit electronic submissions to government agencies.[9]

In the Hans Christian Andersen fairy tale, charlatans deceive the emperor and his advisors into paying for clothing that simply does not exist by claiming that anyone who cannot see the clothing is unfit for his job. When the emperor finally walks down the street displaying what he believes are his new clothes, a child points out his nakedness. The credibility of the innocent child finally cuts through the duplicity and fear of the adults who were afraid to say what they saw and end the charade.

The story of how digital signatures came to be over-hyped and underutilized in electronic commerce is a bit more complex than this fairy tale. In general, digital signatures and public key infrastructures are important examples of cryptography technologies that today play a major role in electronic commerce and information system security. It seems likely, moreover, that the role of cryptography technologies in general and digital signatures and public key infrastructures in particular will continue to grow in the future. So the idea that digital signatures are or will be an important element of Internet commerce is not per se a fraud or an illusion. The specific application of asymmetric cryptography to create the functional analog of an old fashioned manual signature on a contract may well prove to be an illusion, however.[10] There is mounting evidence that trying to use asymmetric

[9] One counterexample to the claim that no Internet contracts are being formed in reliance on digital signature certificates can be found in the electric utility industry, where digital signature certificates are used to identify parties entitled to reserve transmission capacity. This is a result of the Open Access Same Time Information System (OASIS) established by the Federal Energy Regulatory Commission in 1996. *See generally*, Alexander Cavalli and Jane K. Winn, *Internet Security in the Electric Utility Industry*, 38 JURIMETRICS J. 459 (1998); California Independent System Operator Bidder's Policy and Procedures Guide, *available at* http://www.caiso.com/docs/09003a6080/09/e3/09003a608009e33c.pdf (explaining that digital signature certificates are required to participate in an auction of "firm transmission rights" (FTR)). Reservation of transmission capacity within a regulated industry following procedures mandated by federal regulators creates a contract, but this application of digital signature technology does not involve the formation of contracts in reliance on digital signatures and certificates in open Internet commerce because it is product of a regulatory mandate applicable to a closed system.

[10] *See generally*, Ellison and Schneier, *supra* note 2; Roger Clarke, *Conventional Public Key Infrastructure: An Artefact Ill-Fitted to the Needs of the Information Society*, *at* http://www.anu.edu.au/people/Roger.Clarke/II/ PKIMisFit.html; Matt Blaze et. al., *The Role of Trust Management in Distributed Systems Security*, in SECURE INTERNET PROGRAMMING: SECURITY ISSUES FOR MOBILE AND DISTRIBUTED OBJECTS (Vitek and Jensen, eds., 1999), available at http://www.crypto.com/papers/trustmgt.pdf; Dan Geer, *Risk Management is Where the Money is*, *available at* http://www.atstake.com/security/risk_management.pdf.

cryptography as a signature on a contract is like trying to fit a square peg into a round hole, and the effort to get that square peg into that round hole has created a phenomenal sink hole into which countless individuals and organizations have poured vast resources with few tangible payoffs in sight.

Those promoting digital signatures and public key infrastructures have not generally been charlatans of the type Andersen describes, although most may have had pecuniary motives for promoting a particular technology as the "next big thing" in Internet commerce. Since countless individuals and organizations with pecuniary motives routinely promote particular technologies as the "next big thing" in electronic commerce, that is not evidence of bad faith. Rather, promotion of proprietary technologies as supposedly essential elements of the architecture of electronic commerce is business as usual in information economy markets where vaporware and hype are standard operating procedures and parties are routinely locked in mortal combat trying to secure "first mover" advantages. If relatively few technologies have a chance to become incorporated into the network architecture of electronic commerce, but those few that succeed have a shot at vast profits secured by strong network effects, then astute buyers should merely discount such claims accordingly. One of the most interesting puzzles surrounding digital signatures is how so many individuals and organizations that should have known better could have been duped into falling for the hype for so long in the face of mounting evidence of its inaccuracy.

The fear of the bureaucrats in Andersen's fairy tale may have a counterpart in the story of digital signatures hype. In the face of an apparent global consensus that digital signatures would indeed be the "next big thing," those who expressed skepticism about the inevitability of the adoption of this technology risked looking like Luddites[11] or ignoramuses. The global consensus about the inevitability of digital signatures may have at least a partial basis in fact: it is quite likely that this technology will be widely deployed to enhance network security. That outcome remains possible even if it is never used as the analog of a manual signature in traditional contracting practices. The durability of the hype surrounding digital signatures seems also to be due in part to the willingness of individuals to accept at face value in

[11] Luddites were weavers whose trade was being destroyed by mechanized textile mills in England in the late 18th century. Encyclopedia Brittanica Online, Luddites, *at* http://www.britannica.com/bcom/eb/article/0/0,5716,50450+1+49263,00.html?query=luddite (last visited Feb. 14, 2001).

formation they have obtained from questionable sources and repeat it without bothering to confirm the accuracy of factual allegations.

The truth of the factual allegation that digital signatures are the "most popular" form of online authentication in electronic commerce is surprisingly difficult to establish. By all accounts from disinterested parties, it may be one of the least popular forms of online authentication if the standard is the number of contracts formed or the dollar value of transactions entered into in reliance on a digital signature certificate.[12] The simple fact that no one is using digital signatures as signatures to form contracts in open Internet commerce is constantly obscured by references to the fact that pilot projects are underway or have succeeded, or that standards groups are making rapid progress toward completing their work, or that experts all agree that digital signatures are indeed the "next big thing" that no self-respecting electronic commerce cognoscente can live without.

As a result of apparently endless recycling of the contents of public relations press releases[13] or mistaking a description in a statute of a type of business practice for information about the actual popularity of that business practice in the marketplace, the notion that digital signatures are the most widely used form of authentication in electronic commerce today has taken on something of the status of an urban legend. No number of thoughtful refutations of the proposition seem able to kill it off.[14] After it has been defeated in one arena, such as the U.S. Congress, then like the hydra it reappears in its original form and multiplies in new arenas, such as the UNCITRAL working group on electronic commerce[15] or the E.U. Electronic Signatures Directive.[16]

A major part of the problem lies in equating what asymmetric cryptography and a public key infrastructure do in the online context with what a manual signature does in traditional contracting contexts.[17] Traditional signatures play a surprisingly nuanced and com-

[12] *See, e.g.*, Heffernan, *supra* note 4; Higgins, *supra* note 4; Hogan, *supra* note 4.

[13] A search of the "wires" database in Lexis Nexis on February 5, 2001 for stories that included a reference to digital signature, pilot and success or succeed turned up more than 60 press releases issued between 1995 and 2001.

[14] See the sources cited in Clarke, *supra* note 10.

[15] For information about UNCITRAL efforts to develop model legislation governing the use of digital signatures, see the UNCITRAL Working Group on Electronic Commerce Web site at http://www.uncitral.org/english/workinggroups/wg_ec/index.htm.

[16] Directive 1999/93/EC of 13 December 1999 on a Community framework for electronic signatures, *available at* http://europa.eu.int/comm/internal_market/en/media /sign/Dir99-93-ecEN.pdf.

[17] For a civil law perspective on these issues, see Babette Aalberts and Simone van der Hof, *Digital Signature Blindness: Analysis of Legislative Approaches to Electronic*

plex role in traditional contracting practices that prove very difficult to map onto online security technology functions. Not all contracts require a signature to be enforceable, and not all signatures evidence a signer's intent to enter into a binding legal relationship. To apply the term "signature" to the processes performed using asymmetric cryptography, X.509 certificates and a public key infrastructure is at best a metaphor and at worse simply misleading. The poor fit between the metaphorical label "signature" and the security functions performed by digital signatures and public key infrastructure is not immediately apparent to casual observers. Many sophisticated observers who noticed the poor fit had a pecuniary motive not to make the mismatch explicit. Add to these information asymmetries and conflicts of interest the froth and manic energy of an Internet driven speculative bubble, and few were interested in hearing the rather long, complicated story of why digital signatures were *not* the "next big thing."

This Article is part of a symposium on the UETA. Given that the UETA takes no position on the merits of digital signature technology at all, an extended discussion of the lack of success to date in the use of digital signatures in electronic contracts might appear to be a digression from the central focus of the symposium. On the contrary, the "technology neutral" perspective taken by the UETA represents a rejection of the approach taken by many countries which implicitly or explicitly promote the use of digital signatures as an analog to a traditional manual signature and as a necessary element in the process of forming electronic contracts. The UETA approach is a more appropriate legislative response to the question of how digital signatures will be used in electronic contracts because it permits business practices to evolve and Internet authentication security technology to develop through the work of standard-setting processes without mandates from legislatures that appear to be unaware of actual market developments. Managing the rights and obligations of the parties through standards and private agreements permits those with knowledge of market conditions to continue to adapt and evolve information security models more rapidly and more rationally than is possible through the cumbersome and inexact process of legislation.

This Article will summarize the original consensus regarding the role of digital signatures in electronic commerce, explain why that consensus was mistaken on many points, describe commercial applications of digital signatures that are gaining market share today and contrast them with the original consensus, and consider the implica-

Authentication, *available at* http://rechten.kub.nl/simone/ds-fr.htm, which reaches similar conclusions regarding market trends and legislation.

tions of a major misperception of market trends for the future of electronic commerce legislation. A brief description of digital signatures and public key infrastructure is included in the appendix to this article.

II. THE ORIGINAL CONSENSUS: DIGITAL SIGNATURE AS SIGNATURE

The first public key cryptographic system[18] was described in 1976 by Whitfield Diffie and Martin Hellman.[19] A short time later, Ronald Rivest, Adi Shamir, and Len Adelman developed another public key system.[20] The great advantage of a public key system is that it permits individuals to use two different but related keys to maintain the confidentiality of their communications. One key, the private key, is kept secret by the owner, while the other key, the public key, can be widely distributed. The two keys are mathematically related, but one of the features of public key cryptography is that it is computationally infeasible to derive one key from knowledge of the other. A system within which public keys are distributed is often referred to as a "public key infrastructure"[21] (PKI) and is designed to lower the costs associated with distributing public keys while minimizing the risks of fraud and error. The most widely known model of a PKI is based on the model of a telephone directory.[22] This model was first advanced by Diffie and Hellman in a paper published in 1976,[23] and expanded with the notion

[18] See appendix for a discussion of the difference between conventional cryptography, which depends on the use of two identical or "symmetric" keys, and public key, or asymmetric key, cryptography, which depends on the use of two separate but related keys.

[19] SIMSON GARFINKEL, PGP: PRETTY GOOD PRIVACY 49 (1995).

[20] *Id.*

[21] In this article, the term "public key infrastructure" is used to mean any system for regulating the distribution of public keys in a networked environment. The term is often associated with specific designs for distributing public keys, such as the system described in the ABA's Digital Signature Guidelines, *supra* note 1. However, the idea of a "web of trust" associated with the use of Pretty Good Privacy (PGP) encryption program might also be described as a "public key infrastructure" since PGP is based on asymmetric cryptography. See GARFINKEL, *supra* note 19, at 213, for a description of PGP and the web of trust.

[22] Joan Feigenbaum, Towards an Infrastructure for Authorization, Position Paper, 3rd USENIX Workshop on Electronic Commerce (September 1998).

[23] Whitfield Diffie and Martin E. Hellman, *New Directions in Cryptography*, IT-22 IEEE TRANSACTIONS ON INFORMATION THEORY 644 (1976), *cited in* Feigenbaum, *supra* note 22.

of "certificates" by a paper published in 1977 by Loren Kohnfelder, then an undergraduate at MIT.[24]

It has been widely assumed for a decade or more that digital signatures used in combination with digital signature certificates distributed by trusted third parties within a public key infrastructure of some description would revolutionize electronic contracting practices.[25] Digital signatures would provide a stable, reliable mechanism for individuals to manifest their intent to be legally bound by the contents of an electronic record and certificates would form a stable, reliable form of online identity card. Individuals would safeguard their private keys, accessing them only under appropriate circumstances to sign electronic records. Digital signature certificates issued and managed by responsible parties would be included with electronic contracting messages to provide counter parties with a quick, simple way to confirm the real world identity of the author of the electronic communication. The original consensus regarding the role of digital signatures in electronic contracting assumed that there would be a migration away from older online authentication systems[26] toward digital signatures administered within a public key infrastructure. Within that consensus there were vigorous debates about how the private key required to create a digital signature should be kept secure and how the public key infrastructure should be designed and administered. Of course, there were also dissenters from the consensus who argued that the gap between the state of the art of private key security and public key infrastructure design on the one hand, and the needs of transacting parties using the Internet or other networked communication systems today, were simply too great to be bridged in the foreseeable future.[27]

One major obstacle to wide scale deployment of digital signatures in electronic contacting systems seems to be the complexity of the business administration systems it purports to replace. In order to use digital signatures as a functional analog of the messy patchwork of systems now used to authenticate the identity and good faith of contracting parties, the policies and hierarchies that make up a public key infrastructure would have to be integrated with other elements of business information systems that are necessary to permit contract

[24] Loren M. Kohnfelder, Towards a Practical Public-Key Cryptosystem (1978) (unpublished B.S. thesis), *cited in* Rohit Khare and Adam Rifkin, *Weaving a Web of Trust*, v. 1.126 (Nov. 30, 1997), *at* http://www.cs.caltech.edu/~adam/ local/trust.html, n.37.

[25] *See supra* Part I.

[26] Jane Kaufman Winn, *Open Systems, Free Markets, and Regulation of Internet Commerce*, 72 TUL. L. REV. 1177, 1184-1188 (1998).

[27] *See supra* Part I.

negotiations and contract formation to be automated. The policies and hierarchies of individual organizations as well as those supporting the public key infrastructure would have to be standardized for automated transaction processing to be possible among parties with no prior business relationship. After nearly a decade of work in this area, the problem seems no closer to resolution than it was five years ago.

There are several problems with the original consensus regarding digital signatures in electronic commerce. One is whether the metaphor "signature" is appropriate for what can be accomplished with asymmetric cryptography and a public key infrastructure based on certificates. A second is identifying the function a signature serves in traditional contracting practices. A third set of problems are those created by borrowing concepts that make sense in technological standards and trying to insert them into legal analyses in order to change the law applicable to the technology, or borrowing legal concepts and trying to insert them in technological standards in an attempt to expand the range of functions the technology can accommodate.

A. Does the Metaphor of "Signature" Make Sense for Asymmetric Cryptography and Public Key Infrastructures?

The standard model of digital signatures and public key infrastructure is based on the X.509 standard established by the International Telecommunications Union ("ITU").[28] The X.500 standard was developed to facilitate the use of telephone directories over a distributed telephone network such as might be found within a multinational corporation. Different parts of the directory could be stored at different locations on the network, such as the branch office where the individuals whose telephone numbers were listed were employed. One individual wishing to look up a listing for another individual would be able to access the information without regard to where the listing was actually maintained and stored.[29]

When the X.500 standard was being developed during the 1980s by the ITU, the possible use of certificates issued to associate a real world identity with a particular private key was one of the issues ad-

[28] INFORMATION SECURITY COMMITTEE, SECTION OF SCIENCE & TECHNOLOGY AMERICAN BAR ASSOCIATION, DIGITAL SIGNATURE: LEGAL INFRASTRUCTURE FOR CERTIFICATION AUTHORITIES AND SECURE ELECTRONIC 18 (1996). The ITU X.500 series of technical standards provides the basis for constructing a multipurpose distributed directory service by interconnecting computer systems belonging to service providers, governments, and private organizations, on a potentially global scale. WARICK FORD & MICHAEL BAUM, SECURE ELECTRONIC COMMERCE: BUILDING THE INFRASTRUCTURE FOR DIGITAL SIGNATURES AND ENCRYPTION 213 (1997).

[29] It is not clear that the X.509 standard works for telephone directories, but that issue is beyond the scope of this Article.

dressed.[30] The X.509 standard sets forth a description of how a digital signature certificate should be organized. By standardizing the content and presentation of the information contained in a certificate, automated processing of certificates would be possible, as well as exchanges of certificates from different domains. Within a few years, the original X.509 standard, which was designed with a distributed telephone directory in mind, was deemed to be too limited in scope to meet the needs of engineers designing network communication systems and was revised. The X.509 standard that is widely used in electronic commerce applications is version 3 ("X.509 v.3").[31]

The X.509 v.3 standard permits not just an identity to be specified in a certificate, but also policies that govern the certificate's use to be specified. This extension of the X.509 standard to include more than a simple real world identity to include policies that might describe the scope of authorized actions in the online environment was thought to be key to extending the use of digital signature certificates into electronic contracting. For example, an X.509 v.3 certificate might limit its use to transactions below a specified dollar amount, or within a specified geographical region, or to a specified product line. If the electronic contracting systems of counterparties standardize their policies regarding authority to form contracts, then a vendor's fulfillment system could review the limitations in a digital signature certificate and without human intervention make a decision whether or not to accept a purchase order submitted by a prospective purchaser.

Just because an X.509 v.3 certificate contains information about the identity of an individual and may also contain information about the authorized scope of the certificate's use or the authorized scope of the individual's actions online, does not mean it is the analog of a signature. A signature is defined by the Restatement (Second) of Contracts as "any symbol made or adopted with an intention, actual or apparent, to authenticate the writing as that of the signer."[32]

The commentary goes on to point out that a signature is not limited to a handwritten ink signature on paper, but may include a thumbprint, impression of a rubber stamp, or arbitrary code.[33] Under appropriate circumstances, the act of affixing a digital signature cer-

[30] Carl Ellison, *What do you need to know about the person with whom you are doing business?*, October 28, 1997, *at* http://world.std.com/~cme/html/congress1.html (last visited Feb. 14, 2001).

[31] International Telecommunication Union ITU-T X.509 Recommendation (06/97) Data Networks and Open System Communications Directory; Information Technology – Open Systems Interconnection – The Directory: Authentication Framework.

[32] RESTATEMENT (SECOND) OF CONTRACTS § 134 (1981).

[33] *Id.* at cmt. 1.

tificate to a message that has been signed by the private key associated with that certificate might actually constitute a signature, but anyone making such a claim would have to be able to establish a connection between the mental state of the individual to be bound and the act of affixing the certificate and digital signature. The magnitude and complexity of the network architecture and information system security operating at each node on the network necessary to make that connection in a reliable, routine manner is one of the major obstacles now impeding the implementation of digital signature technologies.

There are several obvious problems raised by trying to tie an identity described in a digital signature certificate with the intention of the identified party to be bound to the contents of an electronic record. These include whether the correct person has accessed the private key associated with the digital signature being used; and if a person other than the identified person has used the digital signature, how that person was able to gain access without authorization and who should bear responsibility for that unauthorized access. The breach in security may occur at the level of the end user's failure to take reasonable steps to safeguard access to a private key, or it may occur because the software and hardware used to store the private key have not been made reasonably secure. Before a digital signature can be presumed to be an analog to a traditional manual signature, the behavior, attitudes and sophistication of individuals using the technology will have to be analyzed as well as the security characteristics of the entire system within which an individual digital signature is used. At present, due in part to the lack of standardization among implementations and depth of experience with actual use of digital signature technologies as signatures, that information does not yet exist. In addition, while it may be feasible at present to try to develop and enforce such standards of behavior among participants in a "closed" system in which members agree by contract or system rules on the applicable standards, no one has yet found a feasible way to standardize end user conduct in an "open" environment such as Internet transactions between entities with no prior relationship.[34]

[34] Documents such as certificate policies and certificate practice statements attempt to spell out what behavior is expected of end users, and contracts or system rules setting up closed systems create an obligation on the part of end users to conform their behavior to those standards. E-mail from Rick Hornbeck to Jane K. Winn (Feb. 13, 2001) (on file with author). For a discussion of the difficulties involved in drafting certificate policies or certificate practice statements, see Jane K. Winn, *The Hedgehog and the Fox: Distinguishing Public and Private Sector Approaches to Managing Risk for Internet Transactions*, 51 ADMIN. L. REV. 955 (1999).

B. Why Do Signatures Matter in Traditional Contracting Practices?

When parties form agreements that they expect will be given legal effect, a signature may or may not be part of the process of contract formation. A signature is one type of evidence that is used to show that one of the parties intended to enter into a legally binding relationship, but it is not the only type. In some cases, a signature may not even be a necessary piece of evidence. Just what kinds of evidence of the intention of the parties to enter into a binding agreement will be used in any specific transaction will vary according to the context, including the subject matter for the particular transaction, the communications media the parties are using, the course of dealings between the parties, and the normal business practices in the market or industry. In some situations, the law may require a party seeking to enforce its rights to produce a writing signed by the party against whom enforcement is sought, but such requirements are scarcely universal.[35]

However, once the metaphor of signature had seized the imagination of those looking for new commercial applications for digital signature technology, the search for the "law of signatures" began. In light of the characterization of asymmetric cryptography and a public key infrastructure as a "signature," an obvious research problem was to find the existing law of signatures to determine if it would validate the use of this new technology. Such research efforts uncovered surprisingly little "law of signatures" – some references in digests such as AmJur and some discussion in negotiable instruments law treatises of the proof of signatures on negotiable instruments, but no law review articles at all prior to the 1990s.[36]

Finding a reason why "the dog didn't bark" is always a problematic undertaking, but it is possible to conjecture why signatures were largely a non-controversial subject in legal doctrine until very recently. It is possible that the common law of contracts came to accept a signature as part of the proof that should be offered of intent to be bound so many centuries ago, and that the practice has continued for so long with relatively little change, that the topic scarcely seemed worthy of discussion. Under the medieval common law writ system, signatures were irrelevant to the formation of binding obligations in an era when few could read or write. Rather a covenant under seal was the form of action that was used to enforce what in modern terms

[35] JANE K. WINN & BENJAMIN WRIGHT, THE LAW OF ELECTRONIC COMMERCE § 5.03 (4th ed. 2001) (discussing Statute of Frauds issues).

[36] *See* Winn, *supra* note 26, at 1216-1218 (summarizing the common law of signatures).

might be thought of as a contractual obligation.[37] The pleading rules for covenant under seal were highly formalistic: if a person's seal had been used to authenticate a document, the only defense was to deny the fact that it was the defendant's seal; mere unauthorized use of a seal was not exculpatory.[38] Modern contract law grew out of the writ of trespass, not covenant under seal, when the cause of action for trespass on the case in assumpsit permitted enforcement of undertakings that lacked the formality of covenant.[39] The use of the writ of trespass to give common law courts jurisdiction over undertakings that lacked the formalism of covenants occurred in the 14th century.[40] By the 20th century, methods for proving informal agreements were so well established and so uncontroversial that the topic seems not to have merited sustained discussion outside of relatively limited contexts, such as the statute of frauds or evidence law.

When the technological baseline shifted from some form of handwritten signature and some form of paper record to electronic communications media, anyone trying to map the existing law of signatures onto new commercial practices found no lengthy discussions in general terms of the significance of signatures in contract law. The definition of the issue took roughly the following form: (1) as a practical matter, digital signature technology can replace traditional manual signatures in contract practice; (2) businesses will be discouraged from adopting this new technology, however, if contracts formed with digital signatures are not enforceable to the same extent as traditional paper contracts with manual signatures; (3) if a contract is subject to a statute of frauds requirement of a signed writing, and that requirement is interpreted to mean a manual signature on paper, then that will limit the enforceability of contracts signed with digital signatures; (4) so the significance of "signed writing" within the con-

[37] J. H. BAKER, AN INTRODUCTION TO ENGLISH LEGAL HISTORY 360 (3rd ed. 1990).

[38] FREDERICK G. KEMPIN, JR., HISTORICAL INTRODUCTION TO ANGLO-AMERICAN LAW 215 (3rd ed. 1990). This formalism is similar to that of many "digital signature" statutes which create a "presumption" that a signature is that of the owner of the private key that created it. While a presumption is not the same as a liability rule, the lack of any reliable system for demonstrating who had access to a private cryptographic key at any particular time makes such a presumption tantamount to a liability rule. *See* Jane Winn & Carl Ellison, *U.S. Perspectives on Consumer Protection in the Global Economic Marketplace, comment P994312 to the Federal Trade Commission*, (March 26, 1999), *available at* http://www.ftc.gov/bcp/icpw/ comments/revwin~1.htm. For a UK perspective, see Nicholas Bohm et al., *Electronic Commerce: Who Carries the Risk of Fraud?*, J. INFO. L. & TECHNOLOGY (2000), *available at* http://elj.warwick.ac.uk/jilt/00-3/bohm.html (reaching a similar conclusion regarding contract terms used by UK banks in contracts with customers).

[39] KEMPIN, *supra* note 38, at 374.

[40] *See* BAKER, *supra* note 37 at 375.

text of the statute of frauds must be clarified. Over the last decade or so, many attempts have been made to address this issue, although most of the resulting accounts of the role of signatures in contract law were not neutral, disinterested historical studies.[41] Most of these very recent accounts were colored by the conviction that digital signatures were not only the logical and inevitàble successor to manual signatures on paper, but were also superior to traditional signatures for a variety of reasons.

Studies of the role of signatures in contract law undertaken in this context suffer from at least two distorting assumptions. The first assumption is that the legal significance of signatures generally can be understood by generalizing doctrines found within bodies of law that make express reference to signatures, such as negotiable instruments law or the statute of frauds. Second is the belief that current contract practices lack the technological refinement and rigor that will be possible when new, more powerful authentication technologies are used. These distorting assumptions may result in seriously flawed conclusions if the traditional methods of contract formation never relied exclusively or even primarily on authentication of manual signatures. For example, if the contracting parties were in a long-term relational contract,[42] authentication might rely primarily on oral communications over the telephone, or by making reference to information generated over a long-term course of dealing between the parties.[43] Even in contracts between strangers, there may be a lack of formality that leads the parties to rely on information such as telephone or face-to-face conversations, references from friends, advertising and brand image, or even credit report data to ascertain reliability of an expressed intention to form a binding contract. Obtaining a valid signature is merely one element in a larger problem that the contracting parties are trying to solve: the creation of an agreement that is a "legal, valid and binding obligation . . . [that] is enforceable . . . in accordance with its terms."[44] The focus on the common law of sig-

[41] Even a recent account that attempted neutrality on the question of technological successors to manual signatures on paper would nevertheless be biased by the context of the discussion, namely, identifying what necessary functions manual signatures served in contract practices that could now be served better by electronic equivalents to manual signatures. *See, e.g.*, Winn, *supra* note 26.

[42] Ian Macneil, *Contracts: Adjustment of Long-Term Economic Relations Under Classical, Neoclassical and Relational Contract Law*, 72 NW. U. L. REV. 854 (1978).

[43] For example, a bank customer service representative might ask a bank customer to identify the last three deposits into an account before disclosing sensitive information over the phone.

[44] Special Comm. on Legal Opinions in Commercial Transactions, N.Y. County Lawyers' Ass'n, *Legal Opinions to Third Parties: An Easier Path*, 34 BUS. LAW. 1891,

natures as the antecedent to digital signature laws is too narrow, and overlooks the wide range of factors that might be taken into account in assessing the likelihood that a contract formed by traditional means will be enforceable.

C. What Does "Non-Repudiation" Mean?

A digital signature certificate includes information such as the name of the person or entity to which the certificate was issued, and information about policies governing the contexts in which the certificate may be used.[45] One piece of information a digital signature certificate may include is whether the digital signature is "non-repudiable." If the "non-repudiation" variable in the certificate has been activated, then it should be harder for the person identified in the certificate to deny that the electronic record has been "signed" with his or her private signing key.[46]

If a digital signature validated with reference to a certificate in which the non-repudiation variable has been turned on cannot be denied by the signer, then an electronic contract formed by affixing a digital signature to an electronic record containing a statement of the terms of the agreement should create an obligation that is "legal, valid and binding," and enforceable according to its terms.[47] But flipping on a switch in a digital signature certificate is only one of the many pieces of evidence a court would evaluate before coming to the conclusion that an agreement is enforceable.[48] Notwithstanding this

1914 (1979). Paragraph 4 of the illustrative opinion letter states in full: "The Agreement is a legal, valid and binding obligation of the Corporation and is enforceable against the Corporation in accordance with the terms of the Agreement, except as may be limited by bankruptcy, insolvency, or other similar laws affecting the enforcement of creditors' rights in general. The enforceability of the Corporation's obligations under the Agreement is subject to general principles of equity (regardless of whether such enforceability is considered in a proceeding in equity or at law)." *Id.*

[45] FORD & BAUM, *supra* note 28, at 227.

[46] The X.509 v.3 standard paragraph 12.2.2.3 defines "key usage fields;" one of which is a space for the "non-repudiation" bit. This bit can be used "[f]or verifying digital signatures used in providing a non-repudiation service which protects against the signing entity falsely denying some action." ITU-T Recommendation X.509 ¶ 12.2.2.3 (Aug. 1997); Scott Renfro, *Thoughts on Non-Repudiation*, Feb. 23, 2001, *available at* http://www.renfro.org/scott/writing/non-repudiation-thoughts.txt.

[47] Carl Ellison stated: "The idea that digital signatures could enable electronic commerce through what has come to be known as non-repudiation was first proposed by Diffie and Hellman in their seminal paper, 'New Directions in Cryptography.' The idea has since gained much popularity." Carl Ellison, SPKI/SDSI Certificates, *available at* http://world.std.com/~cme/html/spki.html. Ellison defines "non-repudiation" as "the notion that the keyholder is legally liable for any statement digitally signed by that keyholder's signature key." *Id.*

[48] In addition, "putting the bit in the certificate rather than in the signature itself is sure to complicate expectations. Without lots of education, most people would use

non-congruence between the concept of an enforceable contract and activating the non-repudiation bit in a digital signature certificate, the concept of "non-repudiation" has been creeping into the discussion of electronic contract formation. Muddying the distinction between a legal conclusion and a technological function has contributed to the persistence of the notion that digital signatures are the "next big thing" in electronic contracting.

In principle, it is easy to understand what problems the non-repudiation bit is designed to solve. For example, anyone would understand the difference in meaning between initialing a telephone message taken for another person and signing a mortgage note; between waving a hand to catch the attention of a waiter and waving a hand to make a bid at an auction house; or between shaking hands to greet someone just introduced by a third party, or shaking hands to indicate that a deal has been struck. In the online environment, communications are stripped of many of the contextual clues that help the parties to gauge each other's intentions. The non-repudiation bit could provide an unmistakable signal of intent to form a binding agreement. The problem with conflating the activation of the non-repudiation bit with the formation of a binding contract generally is that it is possible that the bit has been activated without the conscious participation of the party who would be bound by it. If a connection cannot be established between the activation of the non-repudiation bit and the intent of a person capable of forming a contract, then the digital signature certificate is no more effective with the non-repudiation bit activated than with it turned off. Trying to insert the notion of "non-repudiation" into the common law of contracts is at best redundant and at worst misleading.

The term "non-repudiation" as it is used in the X.509 standard is not a term that currently has any significance in contract law. The term "non-repudiation" appears occasionally in other bodies of law, but never with quite the same meaning ascribed to it in the X.509 standard. Even if the notion of "non-repudiation" makes sense in that technical standard,[49] there is no indication that it is a concept that contract law needs to assimilate to retain its relevance in the 21st century. The term has been used in the context of "non-repudiation" of

the same certs for both trivial and big-time-important transactions." E-mail from Neal McBurnett to Jane K. Winn (Feb. 13, 2001) (on file with author).

[49] It may not make sense even in the X.509 standard, however: "The PKI community has . . . repeatedly describe[d] non-repudiation as . . . 'a ...service . . . [that] protects against the signing entity falsely denying some action.'" As anyone with children knows, you cannot prevent someone from "falsely denying" an action. Renfro, *supra* note 46.

collective bargaining agreements under the National Labor Relations Act;[50] "non-repudiation" of an earlier decision by the Atomic Energy Commission;[51] "non-repudiation" of an ERISA plan;[52] "non-repudiation" of a confession by a criminal;[53] "non-repudiation" by a trustee of a fiduciary duty to a beneficiary;[54] and non-repudiation of an agent's act by principal who accepts benefit.[55] The first time the term was used in the context of cryptographic functions, it appeared in the recent *Bernstein v. Department of State* case. However, that case dealt with the issue of whether cryptographic communications were protected speech for First Amendment purposes, not contract formation.[56]

Electronic commerce law is not without any recognition of the importance of clear attribution rules that facilitate enforcement of agreements entered into in reliance on electronic authentication procedures. In UCC Article 4A, the use of a "commercially reasonable security procedure" makes it very difficult for the party from whose system a payment order originated to repudiate an order sent over an electronic funds transfer system, but the term "non-repudiation" is not used in the statute.[57] That legal result follows not simply from one party having used a specific technology but from a prior agreement of the parties to the transaction as well as the objective characteristics of the technology in light of the context of the transaction.

Contract law does recognize the concept of "anticipatory repudiation,"[58] which is a quite different concept than that evoked by the term "non-repudiation." The doctrine of anticipatory repudiation applies to situations where one party to a contract has indicated, by express statement or by conduct, its intention not to perform its obligations under the contract. The other party is permitted a range of options in responding: it may wait to see if performance will ultimately be forthcoming, it may resort to any remedy that would be available in the event of a breach, or it may suspend its own performance. Repudiation here refers to a manifestation of one party's intention not to perform under a contract, not a denial of the existence of the contract.

[50] C.E.K. Industrial Mechanical Contractors, Inc. v. National Labor Relations Board, 921 F.2d 350 (1990).

[51] S&E Contractors v. US, 406 U.S. 1, 62 (1972).

[52] Lemanski v. Lenox Savings Bank, 1996 U.S. Dist. LEXIS 6471 D. Mass. 1996).

[53] U.S. *ex rel.* Johnson v. Lane, 639 F. Supp. 260 (N.D. Ill. 1986).

[54] Norris v. Wirtz, 1985 U.S. Dist. LEXIS 13227 (N.D. Ill. 1985).

[55] Enos v. St. Paul Fire & Marine Ins. Co, 57 N.W. 919 (S.D. 1894).

[56] Bernstein v. U.S. Dept. of State, 974 F. Supp. 1288 (N.D. Cal. 1997).

[57] U.C.C. §§ 4A-201 to 203.

[58] U.C.C. § 2-610; JAMES J. WHITE & ROBERT S. SUMMERS, UNIFORM COMMERCIAL CODE § 6-2 (5th ed. 2000).

Notwithstanding the limited practical relevance of the non-repudiation bit within a public key infrastructure, the concept of "non-repudiation" as a new element of electronic contract law has taken on a life of its own. Take, for example, this discussion of legislation that grants special legal recognition to digital signatures. It conflates "non-repudiation" as a function of technical processes with enforceability as a function of legal processes:

> Illinois . . . distinguishes between an electronic signature (and says an electronic signature is just as good as the autograph) and a digital signature. There are thus two levels of signatures that are contemplated by the Illinois law: the difference between them is that when you use a digital signature (the encrypted kind), you get something called "non-repudiation."
>
> In the context of an autograph, or in the context of an electronic signature, the burden of proof is on the proponent of the admission of the signature into evidence. The question in court is whether the proponent can demonstrate that this is in fact Neil Bardack's "signature" on the document, and that he "authored" the e-mail. The new Illinois law says that if you use a digital signature (the encrypted kind), you can shift the burden of proof to the recipient of the message. With the shift of the burden of proof it is the recipient's responsibility to prove "it wasn't me." This is one of the meanings of non-repudiation.[59]

This discussion seems to assume that, without regard to whether the non-repudiation variable is activated, any digital signature validated with a certificate cannot be "falsely denied" by its putative maker. It also seems to conflate the legal result that the putative signer will not as a practical matter be able to challenge enforcement of the digitally signed record, which is a result dictated by the statute that grants special recognition to digital signatures, with the greater certainty regarding who actually digitally signed the record and with what intent, which is supposed to be a practical consequence of using digital signature technology within a public key infrastructure. This kind of confusion about the practical consequences of using digital signature technology and the enforceability of electronic contracts is all too common in discussions of digital signatures and their relevance to contract law.

[59] Richard Allan Horning, *Legal Recognition of Digital Signatures: A Global Status Report*, 22 HASTINGS COMM. & ENT. L.J. 191, 197 (2000).

Merely confirming that a digital signature can be validated with reference to a certificate cannot take the place of designing a secure system within which electronic agreements can be negotiated and executed. Any form of computer security can be understood as a chain that binds the participants in the information system. The security of the system is only as strong as the weakest link in the chain.[60] The activation of a non-repudiation bit communicates nothing if there is a weak link in the security technology chain that purports to bind the identify of a person to the contents of a digital signature certificate, or the intent of the signer manifested by the act of signing to the concept of non-repudiation. Such a weak link might arise as a result of a confusing interface design which leads individuals to activate the non-repudiation bit without knowing what significance others assign to it; a software application that activates the non-repudiation bit without seeking any confirmation from the person whose intention it purports to express that it should be activated; or a flaw in the design of a security system which permits one person to activate the non-repudiation bit in the digital signature certificate of another person without authorization.[61]

If there is a design flaw somewhere in the public key infrastructure, within which digital signature certificates are distributed and used, then the connection between a person's manifestation of intent to be legally bound by digitally signing a record and the relying party's ability to validate a digital signature with a certificate will be broken and the apparent force of contracts formed within the public key infrastructure will be illusory. The strength of security functions elsewhere in the system may be simply irrelevant in trying to determine the reliability of the system overall. This is why any discussion of how many years it would take to break the security of a cryptographic system by using a brute force attack to guess the value of the key used,[62] is usually a red herring that simply distracts attention from more important issues.

[60] Ellison and Schneier, *supra* note 2.

[61] E-mail from Ben Laurie to Jane K. Winn (Feb. 13, 2001) (on file with author).

[62] *E.g.*, *"Refined Standards, New Concepts Taking Shape,"* EWEEK, Dec. 4, 2000 at 103, *available at* Lexis News ("A code-breaking scheme that takes only 1 second to defeat today's DES [Digital Encryption Standard] would need 149 trillion years to crack a 128-bit implementation of the forthcoming AES [Advanced Encryption Standard."); *cf.*, Bruce Schneier, *Security Pitfalls in Cryptography*, *at* http://www.counterpane.com/pitfalls.html ("Magazine articles like to describe cryptography products in terms of algorithms and key length. Algorithms make good sound bites: they can be explained in a few words and they're easy to compare with one another: '128-bit keys mean good security,' 'Triple-DES means good security,' '40-bit keys mean weak security,' or '2048-bit RSA is better than 1024-bit RSA.' But reality isn't that simple. Longer keys don't always mean more security.").

There are not yet any clear standards regarding what steps users can reasonably be expected to take to keep private keys secure, or how users should be alerted to different possible meanings that may be assigned to the use of a digital signature certificate. If a private key used to make a digital signature is stored on the hard drive of a personal computer and can be accessed by typing in a user ID and password, then the private key is no more secure than the user ID and password. If the user tapes his or her user ID and password to the monitor of the personal computer, it would not be possible to say who had accessed the digital signature. In the absence of well established standards to evaluate the reasonableness of user behavior and human-computer interface designs, the connection between the intention of an individual to be bound by an act executed by computer and the evidence that the act was executed will remain difficult to establish. The fact that a non-repudiation bit was activated in a digital signature certificate will be one piece of information relevant to a determination that an online contract was formed. However, it is only one of many, and hardly sufficient in and of itself to establish that a legal, valid and binding obligation was formed.

III. COMMERCIAL APPLICATIONS OF DIGITAL SIGNATURE TECHNOLOGY

The volume of transactions being conducted over the Internet continues to grow rapidly notwithstanding the lack of acceptance of digital signature technology to create the equivalent of a traditional signature. In the context of business-to-consumer transactions, businesses are relying on the infrastructure that was developed to support mail/telephone order transactions using credit cards.[63] This infrastructure includes verifying certain information such as the billing address before seeing the authorization from the credit card issuer, and running fraud detection software to identify transactions with a higher than acceptable likelihood of being fraudulent. In business-to-business transactions, businesses are relying on a modified version of the old EDI trading partner agreement to enroll parties in a closed system.[64] The trading partner agreement will specify what technology the parties have chosen to identify themselves online and assign responsibility for fraud or error losses that may occur due to a failure in that technology or the failure of one party to implement it correctly.

[63] Jane Winn, *Clash of the Titans: Regulating the Competition between Established and Emerging Electronic Payment Systems*, 14 BERK. TECH. L. J. 675 (1999).

[64] *See generally*, ABA Electronic Messaging Services Task Force, *supra* note 6.

Just because asymmetric cryptography used in a public key infrastructure is not a viable substitute for a traditional signature does not mean that it is not a powerful and important security technology in wide use today. One of the great commercial successes of digital signatures today is the Secure Sockets Layer (SSL) communication security.[65] Part of the key to the success of SSL in the marketplace seems to be that it does not perform any functions analogous to a "signature." It merely permits communications between a browser running on a personal computer and a server to be encrypted in transit, guaranteeing the confidentiality of the communications between the personal computer and the server.[66]

SSL provides some assurance to individuals visiting web sites on the Internet that the sites are genuine merchant sites, and are not operated by a mere hacker masquerading as a legitimate business. The SSL service also provides assurance that transfers of information between the local computer (or "client") and the server are confidential and are received intact. Web server applications that support electronic commerce come with software that manages the keys and the encryption processes in a way that is "transparent" to the visitor to the web site. In Netscape Navigator or Microsoft Explorer, for example, the local user is only alerted to the fact that communications between the client and the server are encrypted when an icon such as a key or a padlock changes, or a dialog box pops up to inform the user that a secure session will be initiated. When an electronic commerce site is set up on the server, public and private keys are generated by a security program, and the public key is used to obtain a certificate from a certificate authority.[67] SSL server certificates are transferred to the client computer for use in the user's browser, either when the browser is first installed on the local client, or in a communication with the server.[68] When a user accesses a Web site that is SSL-enabled, the server first sends a signed copy of the server's digital signature certificate, which the local client verifies. The local client

[65] There is an Internet Engineering Task Force (IETF) standard called Transport Layer Security (TLS) that is based on SSL. TLS Protocol Version 1.0 is available at http://www.ietf.org/rfc/rfc2246.txt?number=2246.

[66] Even that guarantee is not absolute, however. "Handoffs can take place that mean that the certificate validates a different site than what appears in the 'Location:' box in the browser." E-mail from Neal McBurnett to Jane K. Winn (Feb. 13, 2001) (on file with author).

[67] For a more detailed explanation of this process, *see* SIMSON GARFINKEL, WEB SECURITY AND JUSTICE (1997). The role played by certificate authorities in public key infrastructures is discussed in the Appendix.

[68] In fact, several public key certificates are included in the initial installation of recent releases of Netscape's browser. These certificates can be viewed by choosing Security Preferences from the Options pull-down menu in any recent release of Netscape Navigator.

next generates a session key[69] that it encrypts with the server's public key and sends back to the server. All subsequent messages sent between the local user and the server will be encrypted with the session key, so credit card information or other sensitive information cannot be misappropriated even in the unlikely event it is intercepted.

If the metaphor of signature were imposed on the function of SSL, the best that could be said is that the server has a digital signature certificate, but the public key contained in the certificate is used to encrypt something, not to sign something. Even if it was used to sign something, the signature would be of the server, not of the corporation or individual that owned the server. It is hard to imagine under what circumstances a piece of machinery such as a server could be deemed to be party to a contract. Furthermore, there would be no way to show that the user operating the browser software on the personal computer had made a conscious decision to accept something signed by the server, since the authentication of the server's digital signature certificate is made possible through the use of certificate authority certificates that come "pre-installed" in the user's browser software. Given that the user made no decision to trust the certificates preinstalled in the browser software, any act taken following authentication of a digital signature certificate using those pre-installed certificates cannot be said to be taken in reliance on the authentication process performed by the browser software. So if the SSL application creates anything like a "signature," it would be the signature of a piece of machinery reviewed and accepted by a piece of software under conditions that do not permit either the machine or the software to be treated as the electronic agent of either machine owner or the software owner.

Just because asymmetric cryptography has not yet successfully been used in a "signature" application in electronic commerce in the United States does not mean it never will be, however. It is possible that standards for the implementation of digital signatures within a public key infrastructure are now being developed and tested, and will be deployed successfully in the next generation of electronic commerce technologies. There are at least two possible strategies that might make it possible for digital signatures to gain widespread acceptance: the issuance of digital signature certificates by trusted third parties who are prepared to guarantee the accuracy of the contents of digital signature certificates, and a workable system of cross-certification that would permit certificates issued within different

[69] The symmetric key might be based on the Data Encryption Standard (DES) or other recognized standard.

"closed" systems to be accepted by individuals or organizations outside the issuing system. If a trusted third party were willing in effect to guarantee the enforceability of transactions executed in reliance on the certificates, then digital signature certificates would have an obvious value to prospective online trading partners that have no prior relationship with each other. At present, no one has yet found a viable business model for issuing certificates and guaranteeing the contents of those certificates, but this problem may be solved at some point. Cross-certification might be based on a closed system such as a corporation that issues identity certificates to its employees and permits employees to gain access to resources or perform actions within the system based on the information contained within the certificate. In order for the second corporation to accept the first corporation's certificates in making decisions whether to grant access to its own resources or permit actions to be taken by employees of the first corporation, the two corporations will have to standardize many internal policies and procedures. At present, that degree of standardization of corporate policies and procedures has not yet been achieved, but it remains possible that it will be at some point in the future.

One example of a more complex business model for harmonizing corporate policies and digital signature technology sufficiently to permit digital signatures to become an essential element in electronic contract formation is Identrus.[70] Identrus is a joint venture of major banks that are trying to harmonize the traditional risk management functions performed by banks in connection with extending credit to their customers with the demands of Internet commerce for a more powerful system of online identities. The Identrus organization will provide the root certificate authority service for its member banks, which in turn will certify the online identities of bank customers. This service is highly reminiscent of more traditional bank services, such as letter of credit, in which two parties with no prior relationship or other basis for trust ask bank intermediaries to guarantee essential elements of each party's performance under the contract. With the reputation and credit of the intermediary banks to support the reputation and credit of the bank customer's transactions, that would otherwise be declined as too risky, these transactions can now go forward. Individual banks participating in the Identrus system would leverage their existing knowledge of their clients' identity and creditworthiness to guarantee elements of their clients' performance of online contracts.

[70] Information about Identrus is available from the Identrus Web site at www.identrus.com.

While the business model Identrus is based upon is very plausible, it will be sometime before it is clear whether this model is in fact viable in the marketplace. Banks are uniquely positioned to guarantee online identities and creditworthiness because banks have internal expertise in network security associated with electronic funds transfer transactions, and they have access to internally generated information about their customers' business operations. Nevertheless Identrus may still fail to achieve widespread acceptance in the marketplace if it proves more difficult than expected to persuade banks to add online authentication services to the packages of services banks already offer their customers, or if bank customers are not interested in subscribing to such a service at a price that covers the bank's cost of providing such a service.

IV. LAW REFORM AND AUTHENTICATION IN ELECTRONIC COMMERCE

> Never try to teach a pig to sing. It wastes your time and it annoys the pig.[71]

The Uniform Electronic Transactions Act sensibly refrained from trying to teach any pigs to sing when it adopted a "technology neutral" perspective to the formation of electronic contracts. Laws such as the Utah Digital Signature Act, which describe a specific implementation of asymmetric cryptography within a public key infrastructure, have been consigned to the margins of electronic commerce when the marketplace failed to embrace their vision of digital signatures. Merely because a statute does not refer to a particular computer security technology does not mean that the security technology is not vitally important to electronic commerce. Silence within a statute with regard to technological specifics may rather indicate a decision to leave decisions about the network architecture of electronic commerce to private agreements among the parties and technological standard developing organizations. Furthermore, silence within a statute with regard to technological specifics does not imply that the statute does not allocate responsibility among the participants to an electronic transaction for the adequacy of the security systems they adopt.

The two most important provisions in the UETA that have the effect of allocating responsibility among participants to an electronic

[71] American proverb, *cited in* Alice M. Batchelder, *Judges on Judging: Some Brief Reflections of a Circuit Judge*, 54 OHIO ST. L.J. 1453, 1460 (1993). This saying may have originated with Robert A. Heinlein, Time Enough for Love (1974). E-mail from Greg Rose to Jane K. Winn (Feb. 17, 2001) (on file with author).

transaction for the adequacy of the security systems they adopt are section 5(b) which provides that the UETA applies only to transactions in which the parties have each agreed to the use of electronic media;[72] and section 9(a) which provides that an electronic record or signature is attributable to a person only if it is in fact produced by an act of that person. Because the UETA does not contain any presumptions that shift the burden of proof, a person seeking enforcement of rights under a contract executed with electronic media, with the intent to rely on the general validation of such transactions provided by the UETA, will have to prove the other party's consent to the use of electronic media and the other party's actual use of the electronic media in forming the contract. Because there is not yet in wide use a system that reliably binds a person with online actions, including manifesting assent to the use of electronic media or execution of an electronic signature or writing, the party seeking enforcement will have a very considerable burden of proof to meet as a practical matter. The risk that an agreement will not be enforceable, because the party seeking enforcement could not meet its burden of proof, creates economic incentives for parties that wish to enter into electronic agreements on a regular basis to participate in standard setting efforts, such as the development of a system of rules along the lines of the Visa and MasterCard system rules, or the development of clearing house-type agreements that govern the rights and obligations of parties wishing to enter into electronic contracts.

The UETA approach to dealing with the fact that today there is no widely accepted, strong electronic authentication system in place that can be used in Internet commerce creates a rational risk allocation both for the present and for the future. At present, there is a bewildering array of pilot projects and press releases touting solutions to the problem on strong authentication for electronic contracts, but no clear indication of which way the market will move when eventually some more advanced form of authentication technology becomes the new market standard. In a world of many choices but few widely accepted standards, the issue arises which party should bear the risk that a new method of contracting is more risky than one of the methods of contracts in widespread use today – face-to-face agreement; exchange of faxes; telephone or mail order. The UETA puts the increased risk associated with the new method on whichever party ends up seeking enforcement of the contract. That party will have to absorb the costs of researching alternatives and implementing new technolo-

[72] For a criticism of the notion that the use of electronic media in a transaction should justify separate treatment in commercial law, see Joseph Sommer, *Against Cyberlaw*, 15 BERKELEY TECH. L.J. 1145, 1170-1171 (2000).

gies until more secure alternatives to today's Internet communications become available. As a practical matter, that party is more likely to be a business than a consumer, because as repeat players, businesses stand to reap considerable savings by switching from communications media in use today to more sophisticated alternatives.

While it is not possible to predict the future legal framework of online contract formation with any certainty, the automated teller networks in wide use today in the United States and around the world offer an interesting vision of what the future may hold. ATM networks are secured using various security technologies, many of which rely on advanced cryptographic processes that resemble digital signatures created with asymmetric cryptography and administered within a public key infrastructure. Many of the technological standards that govern those technologies and assure uniformity and interoperability are the product of the American National Standards Institute X.9 Accredited Standards Committee for financial services security standards.[73] Among the parties free to set their rights and obligations by private agreement, such as depository institutions and merchants, those agreements may require participants in the system to conform to those standards. Bank supervisory agencies oversee the participation by regulated financial intermediaries in ATM networks to insure that their risk exposure is kept to acceptable levels within the scope of their respective legislative mandates. Consumer liability for using the ATM network, by contrast, is limited by statutory mandates that force the business parties developing, maintaining and using the network to accept responsibility for the security and reliability of the network. ATM networks have expanded their reach outside the borders of the United States through private agreements with foreign banks, merchants and networks. There is no analog in the law of consumer electronic funds transfers to the kind of technology-specific legislation that has been used to promote the adoption of digital signatures.

PKI systems are under development today that may one day bring a level of security to Internet contracting that resembles the security in use today in the ATM networks. The business need for workable, secure contracting systems is tremendous, and huge investments are being made to try to bring products to market that will meet that need. It is possible that interest in PKI systems will increase if businesses experience losses as a result of difficulties in en-

[73] The work of the ANSI X.9 committee is available from its web site at http://www.x9.org/. The ANSI Web store includes a list of standards used in financial services industry, including many based on encryption technologies. *See* http://webstore.ansi.org/ansidocstore/dept.asp?dept_id=80.

forcing electronic contracts formed using less secure technologies. For example, a business that relies on proving that someone clicked through a particular graphical user interface in order to form the contract may find that it is difficult to meet its burden of proof if the contract in question has a high dollar value and the other party vigorously contests the evidence. [74] In addition, market acceptance of PKI technologies would increase if digital signature technologies are offered to businesses in a form that is nearly as transparent to the contracting parties and as reliable as is the security on the ATM networks, or embedded within sophisticated risk management applications that address not just the problem of authentication of counterparties to transactions, but other traditional risks such as credit risks or risk of being drawn into litigation in a remote or hostile forum.

V. CONCLUSION

The other day upon the stair, I met a man who wasn't there.
He wasn't there again today – oh, how I wish he'd go away.

[74] In litigation involving click-through contracts to date, there is no indication the party seeking enforcement has been held to a particularly high standard. An argument that click-through interfaces are unlikely to meet the needs of businesses wishing to form contracts online was made in a posting to the ABA Information Security Committee listserv by Hoyt L. Kesterson II, on February 22, 2001:

> [P]oint and click could be acceptable but that if there was a challenge, then one might be forced to explain that the transactions written on the log were processed during an authenticated session and that even though the log records contained messages that could have been generated by anyone, a review of all the code involved would prove that the messages could have only been written by the user who was authenticated by his password. And while, yes, it is true that someone could have modified the logs after the fact, there are procedures in place that blocks outsiders from accomplishing such tasks, and although people in the company could have modified the logs, this company doesn't do that sort of thing. I suggest that it should be easier to demonstrate that a digitally signed message held in the log was in all probability originated by the user. And while it is true that the code in a pc could be buggered in such a way that it did not reflect the originator's intent to sign, it is probably easier to subvert a point-and-click system. Even if one argues that the possibility of problems in the pc are the same for both point-and-click and digital signature, at least in the digital signature system one can focus on the pc - not on the whole distributed system and network. So while point and click might be acceptable for low risk transactions such as buying a book at amazon.com, it may not be adequate for higher risk applications one must also consider those areas where a persistent record of a signature is required. I have a hard time believing that one could easily prove that a 5 year old point-and-click commitment was done properly.

Id.

Ogden Nash

The problem of online authentication is proving more difficult to solve than Internet commerce pioneers anticipated a decade ago. The push toward greater integration of enterprise applications and more robust open network contracting systems is making the problems of designing and implementing strong online authentication technology ever more complex. As a result, notwithstanding the vast sums of money that have been poured into developing and marketing promising potential solutions, the problem today seems nearly as intractable as it was several years ago.

Over the next five or ten years, huge additional quantities of resources will be poured into finding solutions to the problem of secure online authentication. It is very possible that a standard for secure online authentication will be developed that meets the diverse objectives of transacting parties and that can be incorporated into the next generation of electronic commerce technologies. As a result, it is possible that such a standard might become widely adopted and form part of a new platform for electronic contracting technologies that incorporate the Internet as a communications medium.

With so much present uncertainty regarding what standards will ultimately be developed to meet the needs of contracting parties, and which among those standards will achieve widespread market acceptance, it seems clear that electronic commerce legislation should not try to promote the use of a particular technology. The early digital signature statutes did not merely promote a specific technology, they also promoted a specific standard for the use of that technology. Many years and untold millions of dollars later, no major market participants have been able to promote widespread use of that technology based on that standard. Legislators around the world seem unaware of the difference between the projections of future utilization by interested parties and actual use of a technology. Years of experimentation has revealed that digital signatures are poorly suited for use as a substitute for manual signatures. The effort to make a digital signature work like a manual signature has resulted in the widespread misperception of the role of signatures in the formation of binding electronic contracts. This confusion over appropriate uses of this technology and its contribution to contract formation has in turn led to the introduction of extraneous and unhelpful concepts into the discussion of electronic contract formation, such as "non-repudiation," which only serve to obscure further the terms of the discussion.

The UETA is a notable exception to that trend. It incorporates simple, rational risk allocation rules that can accommodate both the lack of a widely accepted standard today for strong authentication and the possible future development of such standards through the work of technical standard developing organization and private agreements and system rules. While legislation is poorly suited to either describing specific applications for electronic commerce technologies or promoting market adoption of specific technologies, it is well suited to providing rational incentives to the parties capable of shaping the architecture of electronic commerce in the future.

VI. APPENDIX: ASYMMETRIC CRYPTOGRAPHY, DIGITAL SIGNATURES AND PUBLIC KEY INFRASTRUCTURE[75]

Cryptographic security techniques permit information to be shared between two remote parties by minimizing the risk that the information will be intercepted by unfriendly parties or surreptitiously modified in transit. The communicating parties first establish a "cipher" that is used to transform a text into a secure form. The original text is called the "plaintext;" the text after cryptography has been applied is known as the "ciphertext."

The process of converting plaintext to ciphertext is a function of the encryption algorithm. In modern cryptography, encryption algorithms are complex mathematical functions incorporated into software that combine the plaintext with a "key" to produce the ciphertext. The key is a long, seemingly random number, the size of which is measured in bits.[76] The unique value of the key causes the encryption algorithm to produce a unique ciphertext; if the plaintext is modified in any respect, the ciphertext will vary. The better able a cryptosystem is to resist attacks, the more secure it is thought to be. Keys in commercial encryption software use 40-bit, 48-bit, 56-bit, 64-bit, and 128-bit keys; the more bits, the stronger the encryption.[77]

[75] The following discussion is based on JANE K. WINN & BENJAMIN WRIGHT, THE LAW OF ELECTRONIC COMMERCE § 1.04. (4th ed. 2001).

[76] The basic unit of information in programming is a bit, or binary digit. Because computer circuits recognize two levels in electronic current, these two levels of current form the basic binary on/off or 0/1 switches used to communicate data in a digital format. A bit is one unit of information. A byte comprises eight bits. Volumes of digital data are measured in bytes, as in kilobytes (KB), which consist of 1024 bytes, or megabytes (MB), which consist of 1,048,576 bytes.

[77] Responding to a $1,000 challenge from RSA Data Security, a 23-year-old U.C. Berkeley graduate, Ian Goldberg, broke a 40-bit key—the most secure data encryption the US government allows for export—in 3½ hours. There are a trillion possible combinations for a 40-bit key. Goldberg broke it by linking 250 workstations and programming them to run all possible combinations at a rate of 100 billion per hour. Sharon Machlis, *RSA Stunt Shows Up Encryption Weakness*, COMPUTER WORLD, February 3, 1997. In June 1997, responding to a $10,000 challenge from RSA Data Security, a loosely organized group of 14,000 volunteers managed to break a 56-bit key after five months of

In conventional or symmetric cryptography, the same key is used to encrypt and decrypt the message.

Asymmetric cryptography uses two different but mathematically related keys. One key is the "public key," which can be distributed widely without regard to confidentiality; the other is the "private key," which must be kept confidential and carefully secured. The public key may be used to encrypt information that may only be decrypted by the private key; the private key may be used to encrypt information that may only be decrypted by the public key. Because the private key cannot be extrapolated from the public key, the public key may be widely distributed without risk to the secrecy of the private key. Encryption with a public key might be useful in sending a message to the holder of the related private key because such a message can only be decrypted and read by the person in possession of the private key. Encryption with a private key may be useful in sending a message from the holder of the private key because anyone who uses the public key to decrypt the message is reassured that it was sent by no one other than the holder of the related private key.

One problem with public key cryptography is that it may be more computationally intensive than some forms of conventional (symmetric key) cryptography, making it impractical to use public key cryptography to encrypt large files. This drawback of public key cryptography can be solved in several ways, including the use of message digests to ensure the integrity (but not the confidentiality) of the transmitted file, and the use of conventional cryptographic session keys to encrypt the file in combination with public key cryptography to transmit the session key securely. Message digests, or hash functions, help solve the practical problems associated with encrypting entire messages. A message digest produced using a "one-way hash function" is a unique mathematical digest of an entire data file. Identical texts run through the hash function will produce the same digest, but even the smallest change in the text will produce a different digest, alerting the recipient to the fact that the integrity of the message has been compromised. If a guarantee of message integrity

work. The group distributed code-breaking software over the Internet and used idle computers around the world to perform the calculations, with the key being found after trying about a quarter of the 72 quadrillion possibilities. Lynda Radosevich, *Hackers Prove 56-bit DES Is Not Enough*, INFOWORLD, June 30, 1997, at 77. RSA Data Security used the fact that the 40-bit and 56-bit keys could be broken in its efforts to block legislation introduced in Congress to require regulation of encryption using 56-bit or stronger keys, and to encourage the Commerce Department to relax export restrictions on stronger forms of encryption.

rather than confidentiality of the message text is all that is required, a message digest can be an effective solution to the security problem.

It is also possible to combine symmetric key cryptography and asymmetric key cryptography to improve communication security while minimizing the demands made on computing resources. In order for this application to be executed, the sender must already be in possession of the recipient's public key, and the recipient must already be in possession of the sender's public key. The secure e-mail application of the sender of the message generates a "session key" or symmetric key for only one use, usually using a well-accepted form of conventional cryptography such as DES or the International Data Encryption Algorithm (IDEA). The e-mail application then encrypts the contents of the message with the session key before encrypting the session key with the recipient's public key and sending both the encrypted message and the encrypted session key. The recipient uses her private key to decrypt the session key and then uses the session key to decrypt the message.

A digital signature consists of using a private key to encrypt a message digest and then affixing the resulting record to the message itself. In this sense, a digital signature is part of a message that indicates the source of the message and signifies that the message has not been altered in transit. In order for a digital signature to function as the equivalent of a traditional manual signature, there must be a reliable, secure system that permits only the authorized signer to access the private key and affix the digital signature to a message. As with the secure e-mail application, the sender and the recipient must have exchanged public keys prior to sending the digitally signed message. For a digital signature to be affixed to a message, first the signer runs the message through the hash function to produce the message digest. The message digest is then encrypted with the signer's private key, and the result is the digital signature which is affixed to the message. Although the text of the message is not confidential, it is now accompanied by a digital signature unique to the message that can be verified only with the use of the signer's public key.

The verification process takes place when the recipient of the message uses the same hash function as the sender to produce a digest of the message independently. The recipient then takes the public key of the sender and decrypts the message digest from the sender. If the two match, the digital signature has been verified. If a digital signature is removed from the message it was intended to authenticate and attached to a different message, or the original message is modified in any way, then the verification will fail.

The reliability of any cryptographic system depends in large part on the reliability of the system for distributing keys. Symmetric key distribution systems are difficult and expensive to manage. For example, a simple, secure system for distributing symmetric keys is to require a face-to-face meeting between the individuals who will use a key to communicate in the future. Reliable key distribution systems for groups with many members in different geographical locations may require travel by couriers or the use of other cumbersome or expensive secure communication systems.

Key distribution problems in asymmetric key cryptography systems may be less difficult to solve than key distribution problems in symmetric key distribution systems because a public key can be widely distributed without fear of compromising the security of the private key. In a symmetric key system, by contrast, the single key must be kept private by both parties and never be distributed widely; if that key ever falls into the hands of an unintended recipient, the parties should stop using it and replace it with a new key. Key management remains an issue with public key cryptography, however, because once the private key has been created and the related public key distributed, the owner of the private key is at risk if the security of the private key is compromised, because an attacker could then impersonate the true owner of the key.

After keys have been distributed, their use must be managed. Private keys must be kept secure and under the exclusive control of the person or object associated with the key and users must be notified whenever the security of a private key is compromised so that the corresponding public key is no longer used. Systems developed to manage keys are referred to as public key infrastructures (PKIs). There are many different approaches to designing a PKIs. Systems that facilitate the verification of digital signatures between strangers over the Internet are usually referred to as "open PKI" solutions. Systems that rely on binding, in advance, all the relevant parties to a digitally signed transaction with a system of contract that spells out the legal consequences of using public key cryptography or that implement a PKI in a bound community with a defined group of members are usually referred to as "closed PKI" solutions.

One solution to the key distribution problem that may lower the costs of maintaining the public key infrastructure is to find a trusted third party to be responsible for binding an individual with a public key.[78] One type of trusted third party is a certification authority (CA).

[78] Other solutions include the "web of trust" used in the Pretty Good Privacy system (PGP) of digital signatures. Individuals indicate their trust in the public keys of other individuals by "certifying" them with their own digital signatures; the PGP program reviews the digital signatures that certify the validity of a new public key to determine if

The CA reviews some evidence that a particular individual is appropriately using a digital signature, and then issues a "certificate" containing a copy of the public key of the individual signed by the CA. The individual seeking certification is known as a "subscriber." Anyone who wishes to verify the digital signature of that individual may use the public key of the individual in the certificate. A person who uses the certificate to verify the digital signature is known as the "relying party." A CA establishes policies that govern the circumstances under which it issues certificates; these policies are then published in a "certification practice statement" disclosing those policies to any potential subscribers or relying parties.

In order for a certificate issued by a particular CA to be acceptable to a prospective relying party, the CA must establish its trustworthiness in some way. That trustworthiness may depend on its reputation in traditional business transactions, or the CA may in turn be a subscriber of a higher CA, and use the certificate of the higher CA to reassure subscribers and relying parties that it is not a bogus CA. The CA at the pinnacle of the CA hierarchy is known as a "root" CA in such a system; a government might provide root CA services to reduce the possibility of rogue CAs.[79]

Another fundamental key management issue to be resolved is how the revocation or termination of keys should be handled once they have been widely distributed. A key owner may wish to revoke a public key if the security of the private key has been compromised, or may have a policy of retiring keys after a certain period of time has passed to reduce the probability of the key being broken in an attack. In addition, the CA may wish to cancel a certificate if it becomes aware of improprieties in its issuance or at the request of the subscriber. A relying party should investigate the current status of a certificate before relying on it to learn if it is still effective. A CA might provide an authorization service like that provided by credit card companies, in which a potential relying party contacts the CA before relying, to learn if the certificate is still outstanding and has not been revoked for any reason. However, if the CA's practice statement limits its review to the time of issuance, then there is no ongoing monitoring by the CA of the subscriber's status. The CA may maintain a "certification revocation list" where notices by subscribers are posted as soon as received, and that any prospective relying party should check before verifying a digital signature.

it has been signed by someone the recipient trusts. *See* SIMSON GARFINKEL, *supra* note 19, at 235.

[79] *See* Winn and Wright, *supra* note 35, § 5.07 (describing recent e-commerce legislation including CAs); *see also*, Baker & McKenzie Web site, Electronic and Digital Signature Resources, *at* http://www.bmck.com/ecomme rce/topic-esignatures.htm.