Legislating Market Winners: Digital Signature Laws and the Electronic Commerce Marketplace

C. BRADFORD BIDDLE*

Abstract: "Legislating Market Winners" argues that certain enacted digital signature laws are premised upon false assumptions, and inappropriately enshrine a business model which would not evolve naturally in the marketplace. In attempting to solve an unsolvable liability allocation problem, such legislation harms consumers and the future evolution of electronic commerce. The article points out that alternative business models can solve the liability allocation problem. Despite obvious flaws, legislation of this type continues to be proposed, partly because the infrastructure created by these laws coincides with the needs of key escrow proponents. Ultimately the article argues that digital signature laws, which impose a

---

* Brad Biddle is the author of several articles on the law and policy of public key cryptography, including *Misplaced Priorities: The Utah Digital Signature Act and Liability Allocation in a Public Key Infrastructure*, 33 SAN DIEGO L. REV. 1143 (1996). He is Vice Chair of the Electronic Commerce Subcommittee of the American Bar Association's Committee on the Law of Commerce in Cyberspace, and is an associate in the San Diego office of Cooley Godward LLP, where he served on the legal team advising the Internet Law and Policy Forum's Working Group on Certification Authority Practices. The opinions expressed in this article are his own individual views. He thanks Eric Schlachter for his thoughtful suggestions concerning this article, and dedicates the article to Connor Thomas Biddle. This article is based upon an article of the same title which appeared originally in the Summer 1997 WORLD WIDE WEB JOURNAL.

particular view of electronic commerce, should be abandoned, in favor of laws which remove specific, well-defined barriers to electronic commerce and which allow the electronic commerce marketplace to evolve unfettered.

I. INTRODUCTION

The argument goes something like this: Internet commerce is hampered by the authentication problem. There is no reliable way to ensure that the sender of an electronic transmission is in fact who they purport to be. Digital signatures, supported by a "public key infrastructure" of certification authorities (CAs) and certificate databases, can solve this authentication problem. CAs will not emerge under the current legal regime, however, because they face uncertain and potentially immense liability exposure. Additionally, the legal status of digitally signed documents is unclear. Therefore, legislation is needed which defines and limits CA liability and which establishes the legality of digitally signed documents. Such legislation will solve the authentication problem and result in robust Internet commerce.

This argument has captured an influential segment of the legal community, and has led to the enactment of "digital signature legislation" in several U.S. states and foreign nations. Unfortunately, the argument is built on fundamentally flawed assumptions and the legislation enacted based upon it is correspondingly flawed. Much (but not all) of the digital signature legislation enacted to date presumes a vision of electronic commerce that simply is not tenable, and which would not "naturally" evolve in the marketplace. This legislation poses the risk of profoundly distorting an infant market and locking in business models which are harmful to consumers and to the future development of electronic commerce.

The type of public key infrastructure (PKI) envisioned by many of the existing digital signature laws is not viable. The problem is liability. Digital signature legislation drafters have assumed that the potential liability exposure faced by CAs is somehow a flaw of the existing legal regime. This is an erroneous assumption: the liability exposure faced by CAs under the "open PKI" model envisioned by legislation drafters is a product of a business model that cannot internalize the costs associated with its implementation. Moreover, in attempting to limit the liability exposure of CAs, current digital signature laws shift an immense liability burden onto consumers who use the infrastructure envisioned by these laws. Putting this type of liability burden on consumers violates long-held tenets of public policy, and is a result which consumers would reject in any truly "bargained for" transaction.

Digital signatures will undoubtedly play a significant role in electronic commerce. However, rather than being implemented in the "open PKI" model envisioned by various digital signature laws, digital signatures are more likely to be utilized under a "closed PKI" model. Under a closed PKI system, the liability problems associated with digital signatures become much more manageable. Closed PKI offers several other advantages as well. This article describes the differences between open and closed PKI, and suggests that, in the absence of legislative displacement, certain marketplace trends indicate that closed PKI is indeed the likely market winner.

The open PKI model can and should compete against closed PKI and other authentication technologies, and should not be accorded special legal status via legislation. Such legislation is unnecessary: the "contractual privity problem" which is used to justify open PKI legislation is a red herring. Commercial CAs utilizing the open PKI model can compete in the marketplace without special PKI legislation. These CAs are unlikely to succeed, not because of flaws with the legal system, but because the open PKI model is not a winning business model.

Despite raising the very peculiar specter of regulating an essentially nonexistent industry (CAs), and despite increased recognition of the problems associated with the very specific vision of electronic commerce embodied in these digital signature laws, laws based on the open PKI model continue to be proposed and implemented. This article suggests that one of several factors behind the continued momentum of this legislation, particularly at the federal and international levels, is its synergy with cryptographic "key escrow" proposals. While digital signature legislation ostensibly addresses the use of cryptography only for the purposes of authentication, and not for confidentiality, the infrastructure created by these laws is ideal for implementing a key escrow scheme.

Ultimately this article argues that digital signature laws which impose a particular view of electronic commerce should be abandoned. Laws which remove specific, well-defined barriers to electronic commerce—such as unnecessary "writing" or handwritten signature requirements—and which allow the electronic commerce marketplace to evolve unfettered should be encouraged.

II. BACKGROUND: DIGITAL SIGNATURES AND PUBLIC KEY CRYPTOGRAPHY

Digital signatures are a particular application of public key cryptography. No attempt will be made here to explain the rather complex underlying technology in any detail; readers who are unfamiliar with basic cryptographic terminology and techniques should consult some of the many excellent sources available which can provide the relevant technical background.[1] The importance of understanding the technology cannot be overstated: at least some of the flaws in cryptography-related legislation can be attributed to inadequate technical knowledge on the part of policymakers. At the risk of oversimplifying to the point of inaccuracy, creating a digital signature involves encrypting a numerical representation of an electronic message with a private encryption key, which the owner keeps secret; verifying a digital signature involves decrypting the encrypted data using a related public encryption key, which can be made widely available.

Lawyers have largely focused on what digital signatures can accomplish, if implemented in a particular ideal setting. If Alice signs an electronic document with a digital signature and sends it over the Internet to Bob, ideally Bob can be assured that, first, the message really came from Alice. Digital signatures can provide assurance that a message has in fact come from its purported sender, a quality called "data origin authentication." Second, Bob could know that the message he received is the exact message that Alice sent. A digital signature enables a recipient of a message to verify that a message has not been intentionally or accidentally altered during transmission, a quality known as "message integrity." Third, Alice cannot later deny that she sent the message. No one else could have sent the message but Alice, and Bob can prove that unequivocally. This quality provided by digital signatures is known as "non-repudiation."[2]

Two difficult problems must be overcome in order to actually fulfill the promise of digital signatures. The first is identification. Alice may

---

1. Good sources include BRUCE SCHNEIER, E-MAIL SECURITY: HOW TO KEEP YOUR ELECTRONIC MESSAGES PRIVATE (1995), which is highly recommended as an excellent general introduction to the fundamentals of cryptography. Another excellent introduction to cryptography and digital signatures is RSA Laboratories, *Answers to Frequently Asked Questions About Today's Cryptography* (visited Jan. 17, 1998) <ftp://ftp.rsa.com/pub/labsfaq/>. A more sophisticated and comprehensive introduction to cryptography can be found in BRUCE SCHNEIER, APPLIED CRYPTOGRAPHY: PROTOCOLS, ALGORITHMS, AND SOURCE CODE IN C (2d ed. 1996).

2. *See, e.g.*, Michael J. Ganley, *Digital Signatures and Their Uses*, 13 COMPUTERS AND SECURITY 385 (1994).

not have sent the message to Bob at all. Instead, a forger may have generated a cryptographic key pair, and entered the public key in a public key database under the name "Alice." "Alice" and Bob may have entered into some business arrangement whereby Bob performed some service for "Alice," and "Alice" promised to pay Bob. When Bob attempts to enforce his electronic contract however, he will find that he has been the victim of fraud. Digital certificates, issued by "trusted third parties" called certification authorities, are one attempt to solve this problem of identification.

Certificates are digitally-signed electronic documents issued by CAs that attest to the connection of a public encryption key to an individual (or other entity). The process might work like this. Alice would generate her public and private key pair. She would then present her public key to a CA, along with some form of identification. The CA would check the identification and take any other steps necessary to assure itself that Alice was indeed who she claimed to be. The CA would then give Alice a certificate attesting to the connection between Alice and her public key.

The CA must also somehow provide assurance that it is bound to its public key, which is used to verify Alice's certificate. Thus, the CA could have its own certificate, signed with the digital signature of a "higher level" certification authority. This higher level certification authority might be (as under some of the enacted digital signature laws) a government agency.

When Bob received a message from Alice signed with Alice's digital signature, he could obtain Alice's certificate either directly from Alice or from an online database. If the signature on the message could be verified using the public key listed in the certificate, and the CA's signature verified as well, ideally Bob would know that a CA had authenticated Alice's identity, and that he was not dealing with someone else posing as Alice.

The second vexing problem presented by public key cryptography is the security of private keys. If a forger somehow discovers Alice's private key, that forger can digitally sign Alice's name on documents. If a criminal discovered a certification authority's private key, that criminal would have the means to commit widespread fraud. As a practical matter, in any large-scale system utilizing public key cryptography some private keys will become compromised, and the certificate containing the corresponding public key will need to be revoked.

Certificate revocation lists (CRLs) are designed to prevent people from relying on a compromised or otherwise revoked public key/private key pair.

A CRL is a list of public keys that have been revoked prior to their expiration date. If the private key is compromised, or the key pair is no longer in use for some other reason, the public key would be placed on a CRL. Thus, before Bob relied on the message that he received from Alice, he would check to make sure that Alice's certificate was not on a CRL.

III. DIGITAL SIGNATURE LEGISLATION

A segment of the legal community, noting the authentication problems associated with the Internet, became increasingly enamored with the possibilities of digital signatures. Beginning in 1992, efforts began in earnest to develop legal rules to support the type of public key infrastructure described above. Many of these efforts took place within the framework of the Information Security Committee of the American Bar Association's Section of Science and Technology (the "Information Security Committee").[3]

A primary assumption of this group of lawyers was that the specter of large, uncertain liability exposure would prevent the emergence of commercial CAs.[4] The liability problem has several aspects. First, if

---

3. General information about the Information Security Committee can be found on the Internet at (visited Jan. 17, 1998) <http://www.abanet.org/scitech/home.html>. *See* DIGITAL SIGNATURE GUIDELINES: LEGAL INFRASTRUCTURE FOR CERTIFICATION AUTHORITIES AND SECURE ELECTRONIC COMMERCE 1 (1996) [hereinafter DIGITAL SIGNATURE GUIDELINES].

4. The effect of this Guideline [3.14] is to preclude liability for breach of a duty not included in these Guidelines. The role of a certification authority is developing, and few will enter this uncharted area of business without first having the basic rules established with sufficient clarity to enable an evaluation of the legal risks of the new business . . . . [T]his Guideline seeks to limit the legal risk to those described in these Guidelines.

DIGITAL SIGNATURE GUIDELINES, *supra* note 3, at 77. The Utah Department of Commerce reported the drafting committee of the Utah Digital Signature Law commenting to the now-enacted amended version of the Utah Digital Signature Act:

[The Act] clarifies the liability and risk of liability that certification authorities bear . . . . [A] certification authority must be able to assess and manage its risk of exposure to possible liability, and one of the principal impediments to the emergence of certification authorities has been the uncertainty of the legal risks such a business would undertake.

DIVISION OF CORPORATIONS AND COMMERCIAL CODE, UTAH DEPARTMENT OF COMMERCE, UTAH DIGITAL SIGNATURE LAW: TECHNICALLY AND LEGALLY SECURE ELECTRONIC COMMERCE 58 (1995) [hereinafter UTAH DIGITAL SIGNATURE LAW]. *See also* A. Michael Froomkin, *The Essential Role of Trusted Third Parties in Electronic Commerce,* 75 OR. L. REV. 49, 109-10 (1996).

a criminal defrauded a CA and induced the CA to issue a false certificate, the criminal could impose losses on a large number of third parties who would rely on the erroneous certificate. The CA could take every reasonable step—or even extremely costly, exceptional steps—to confirm identity, but still issue an erroneous certificate. If every party who relied on the certificate had a claim against the CA for any consequent losses, the CA's potential liability could be staggering. CAs would be forced to go to extraordinary lengths to confirm identity in every situation in order to avoid potential liability exposure, even when parties to a given transaction may have been satisfied with a less rigorously-procured (and, thus, less expensive) certificate.

Additionally, CAs face potential liability for claims by parties who rely on a certificate after the private key associated with the public key listed in the certificate is stolen by a criminal who then creates forged digitally-signed documents. This type of harm would be difficult for CAs to prevent: they have little or no control over the care a "subscriber"[5] takes in protecting their private key from misappropriation. If CAs bear this risk, it will be reflected in the price of certificates, which might then be uneconomically high.

Lastly, CAs face catastrophic liability exposure if their private key is compromised. If a criminal obtained a CA's private key, they could commit widespread fraud. Additionally, once the compromise was discovered, all certificates issued by that CA would have to be revoked and new certificates issued, imposing costs on all the subscribers of that CA. If CAs face liability for these potentially immense losses, entrepreneurs might choose not to enter the CA business at all.

The liability problem was perceived to be particularly intractable because of a "contractual privity problem." CAs could presumably enter into contracts with their subscribers, and allocate risk between the CA and subscriber via contract mechanisms (i.e., the CA could offer certain limited warranties to the subscriber, and limit potential liability to an agreed-upon amount). However, the lawyers looking at this issue believed that CAs typically would not be able to establish a contractual relationship with the parties who would rely on certificates, in order to

5. "Subscriber" is the term frequently applied to the individual or entity which obtains a certificate from a CA. *See, e.g.*, DIGITAL SIGNATURE GUIDELINES, *supra* note 3, at 50-51.

allocate risk by contract.[6] Therefore, these lawyers concluded that legislation was needed which set out the duties of all parties in this public key infrastructure and which allocated liability appropriately.

The Information Security Committee planned to release a "U.S. Model Digital Signature Act" in June of 1995.[7] Increasingly, however, some members of the committee grew dissatisfied with the planned legislative approach. Ultimately, for a variety of reasons, the plan to release model legislation was dropped. In October, 1995, the Information Security Committee did release an Exposure Draft of its *Digital Signature Guidelines*, which it described as "general, abstract statements of principle, intended to serve as long-term, unifying foundations for digital signature law across varying legal settings."[8] The Guidelines, released in their final version in August 1996, set out duties for CAs, subscribers, and relying parties, consistent with the vision for a PKI described above. The *Digital Signature Guidelines* avoid taking positions on certain detailed issues that legislation in this area would address, however.[9]

In collaboration with the Information Security Committee,[10] the state of Utah began developing digital signature legislation, and the Utah Digital Signature Act was enacted (with considerable fanfare) in March

---

6. *See id.* at 19 (noting that "[t]he relationship between a certification authority and subscriber may be primarily contractual . . ." but that "[t]he duties of a certification authority to a third party relying on a certificate are rooted mainly in legal proscriptions against fraud and negligent misrepresentation.").

7. The saga of the "U.S. Model Digital Signature Act" and related developments, summarized in the next four paragraphs of the text, is detailed in, Comment, *Misplaced Priorities: The Utah Digital Signature Act and Liability Allocation in a Public Key Infrastructure*, 33 SAN DIEGO L. REV. 1143, 1164-67 (1996), and in the sources cited therein.

8. DIGITAL SIGNATURE GUIDELINES 20 (Exposure Draft 1995). The Guidelines clearly contemplate serving as a framework for legislation: the final version of the Guidelines make several references to "implementing legislation." DIGITAL SIGNATURE GUIDELINES, *supra* note 3, at 80, 91.

9. *See, e.g.*, DIGITAL SIGNATURE GUIDELINES, *supra* note 3, at 80 (noting that the Guidelines are "intentionally silent" on the duty of care required of holders of private keys).

10. The original, 1995 Utah Act and the [now-enacted] Draft Amended Utah Act were developed in collaboration with the Information Security Committee of the Section of Science and Technology of the American Bar Association (the 'ABA Committee'). The Utah Digital Signature Legislative Facilitation Committee and the ABA Committee shared some common redactive personnel, and the ABA Committee has for the most part taken an active and supportive interest in the Utah Act.

UTAH DIGITAL SIGNATURE LAW, *supra* note 4, at 18. The Information Security Committee endorsed the Utah Act "in principle." Letter from Michael S. Baum, Chair, Information Security Committee, *Resolution by the Information Security Committee on the Proposed Utah Digital Signature Legislation*, Section of Science and Technology, A.B.A. (1994) (on file with author).

of 1995.[11] Under the Utah Act, a government agency assumes the obligations of being a "top level" CA and is charged with policymaking, facilitating implementation of digital signature technology, and providing regulatory oversight of private sector CAs through a comprehensive licensing scheme. Licensing under the Utah Act is voluntary; however, licensed CAs are offered certain legal benefits (primarily limited liability). The Utah Act imposes detailed duties on CAs, subscribers, and relying parties which are consistent with the *Digital Signature Guidelines*, allocates liability among these parties, and accords special legal status to digitally-signed documents created using the services of a licensed CA.[12]

A number of states turned to the Utah Act as model digital signature legislation, a process encouraged by the drafters of the Utah law. In several public communications, a prominent Information Security Committee member who was also involved in the drafting of the Utah Act indicated that the Utah Digital Signature Act was substantively identical to the unreleased ABA Model Digital Signature Act. In the wake of the enactment of the Utah Act, digital signature legislation based on the Utah law was proposed in nearly a dozen states. By mid-1997 Washington and Minnesota had enacted laws which closely tracked Utah's, and similar bills remain pending in many states.[13] The Utah Act proved influential even when not explicitly followed: California considered and then rejected the Utah model, enacting a non-technology-specific bill designed to address transactions with government entities.[14] Early drafts of the regulations, designed to implement the California law, closely followed the Utah model, however.

The "Utah/ABA Guidelines" model has also proven influential at the international level. Malaysia recently enacted legislation based upon the

---

11. The Utah Digital Signature Act was enacted by 1995 Utah S.B. 82, creating UTAH CODE ANN. §§ 46-3-101 to -504 (1993 & Supp. 1997). It was significantly amended by 1996 Utah S.B. 188, which repealed and reenacted large portions of the Act. The Act is found in its amended form at UTAH CODE ANN. §§ 46-3-101 to -504 (1993 & Supp. 1997).

12. The various provisions of the Utah Act are described in detail in Comment, *supra* note 7, at 1153-63.

13. The Chicago law firm McBride Baker and Coles provides an excellent summary of enacted and pending digital signature legislation at (visited Jan. 17, 1998) <http://www.mbc.com/ds_sum.html>. Readers interested in determining the latest legislative developments should refer to that site.

14. CAL. GOV'T CODE § 16.5 (West 1995 & Supp. 1997).

Utah Act;[15] similar legislation is under consideration in Australia, Canada, Germany, Singapore, and the European Union.[16] The United Nations Committee on International Trade Law (UNCITRAL) is also studying Utah-style model legislation.[17]

Recent legislation proposed at the federal level in the United States and in the United Kingdom adopts the Utah/ABA Guidelines model with an added twist: key escrow.[18] Under these proposed laws, CAs (or TTPs—trusted third parties—as they are called under the U.K. bill) not only serve to bind subscribers to their public encryption keys used for authentication purposes, but also serve as key escrow agents, verifying the escrowing of keys used for confidentiality purposes.

Not all legislative bodies have jumped on the Utah/ABA Guidelines bandwagon. Several U.S. states enacted legislation which addressed "electronic signatures" and other non-public key methods of authenticating electronic transmissions. The most thoughtful legislative effort to date has occurred in Massachusetts, where concerns over the market-distorting effects of Utah-style legislation led to a proposed bill which is aimed simply at removing existing legal barriers to electronic commerce.[19] Similarly, the influential National Conference of Commissioners on Uniform State Laws (NCCUSL) is considering a uniform law on electronic contracting and has not been receptive to arguments in favor of a Utah-style law.

IV. OPEN PKI: THE LIABILITY "TRILEMMA"

The Utah Act and its progeny, and the ABA Guidelines, are premised on an "open system" or "open loop" model of a PKI. The open PKI model envisions that subscribers will obtain a single certificate from an independent third-party CA which certifies that subscriber's identity.

---

15. *Bill For Use Of Digital Signatures Approved,* NEW STRAITS TIMES (Malaysia), May 6, 1997, at 8.

16. The McBride Baker and Coles Web site, *supra* note 13, provides a good summary of international developments as well.

17. *UN Trade Law Commission Concludes Session, Adopting Model Law On Cross-Border Insolvency,* M2 PRESSWIRE, June 4, 1997 ("the Commission also continued its work on legal aspects of electronic commerce, including the questions of digital signatures and certification authorities").

18. *See infra* notes 47-49 and accompanying text.

19. The Information Technology Division of the state of Massachusetts has compiled an excellent set of links to PKI related documents, including their own proposed legislation, at (visited Jan. 17, 1998) <http://www.magnet.state.ma.us/itd/legal>.

Certificate holders will then use that certificate to facilitate transactions with potentially numerous merchants and/or other individuals.[20]

As discussed above, the open PKI scenario implicates considerable liability risk. Proponents of the open model, enamored with what digital signatures can potentially accomplish, have attributed this risk to flaws in the existing legal regime that must be addressed legislatively. This conclusion is wrong. The liability exposure faced by CAs under the open PKI model is the product of a business model that cannot internalize the costs of the inevitable fraud that will result under any public key-based system. The resulting liability problem is unlikely to be solved at all in the open PKI model, and certainly cannot be solved with any one-size-fits-all legislative solution.

Here is one aspect of the problem: if criminals can obtain something valuable by expropriating an individual's private key, they will. There is simply no practical way to keep private encryption keys truly secure. Proponents of digital signature technology frequently mention the concept of storing private keys on tamper-proof smart cards. While smart cards, particularly smart cards that incorporated biometric measures designed to prevent unauthorized use of a misappropriated card, would undoubtedly promote the security of private keys, at this point this is simply wishful thinking—this type of smart card is not commercially deployed in any meaningful way, and is unlikely to be in the foreseeable future. There is currently no realistic way to truly secure private keys, and this problem is going to get worse before it gets better.[21]

---

20. The "open PKI" vs. "closed PKI" nomenclature is problematic is some respects, and is not endorsed here with great enthusiasm. An alternative formulation might posit "Generic Identity Certification Authorities" (open PKI) against "Context-Specific Certifiers" or "Context-Specific Certificate Issuers" (closed PKI). Additionally, the terminology "open PKI" and "closed PKI" is not meant to imply open networks or open standards vs. closed networks or proprietary technology, but rather to describe specific business models. Open systems, in the more typical sense of the term, are associated with larger markets, lower barriers to entry, more rapid technological progress, and ultimately all the benefits of a more competitive market structure. This article argues that it is the closed PKI model which can better provide such benefits.

21. Consider the seemingly endless stream of announcements concerning security breaches in the leading commercial Web browsers, some of which would allow a malicious hacker to access files—including potentially a private encryption key—stored on the computer of an unsuspecting user. Many of these problems are cataloged at (visited Jan. 17, 1998) <http://www.cs.purdue.edu/coast/coast.html>.

Private keys will be expropriated, and third parties will rely on ostensibly valid but fraudulent documents and suffer losses. The aggregate losses could be quite sizable, judging from analogous contexts: Mastercard and Visa lose over $1 billion per year to fraud; phone companies claim to lose $3 billion a year to fraud; in the city of Los Angeles alone fraudulent real estate document filing is said to have cost $131 million in a twenty-seven month period.[22] Who will bear these losses? There are three primary choices: 1) the relying party; 2) the individual whose key was used to sign the document; or 3) the CA who performed the initial binding.

Under the Utah Act, the individual whose key was used to sign the document bears unlimited liability if they failed to exercise "reasonable care" to protect their private key.[23] The Act also imposes difficult evidentiary burdens on the individual. So, if a subscriber—named "Grandmom," just to put things in perspective—does not exercise reasonable care, and her key is stolen resulting in losses totaling $25,000 prior to revocation of her key, that subscriber bears the loss—i.e., Grandmom loses her house. Or, if Grandmom *does* exercise reasonable care and her key is still misappropriated, she must present a court with "clear and convincing" evidence (the standard under the Utah Act) to overcome the presumption that a document signed with her digital signature was in fact signed by her. In either case, the result does not comport with well-established consumer protection laws (compare with the legislatively-imposed $50 consumer liability limit for credit card losses, or the fact that one cannot be bound by a fraudulent handwritten signature).[24] Moreover, no rational consumer would agree to accept this level of risk in a marketplace transaction. The benefits of having a certificate simply do not outweigh the very real possibility of facing extraordinarily large unreimbursed losses.

---

22. In 1994 Mastercard reported a loss of $486 million due to credit card fraud; Visa's fraud loss was $645 million. Robert Jennings, *Fraud is Stealing Holiday Joy from Credit Card Companies*, AMERICAN BANKER, Dec. 7, 1995, at 1. Phone companies estimate that they lose about $3 billion to calling card fraud and other types of fraud. Peter Sinton, *Visa Has Sights Set on Credit Card Fraud*, SAN FRANCISCO CHRONICLE, Sept. 14, 1994, at B-1. L.A. County District Attorney Gil Garcetti estimated that county residents were cheated out of $131 million between July 1990 and November 1992. Timothy J. Moroney, *Review of Selected 1995 California Legislation*, 27 PAC. L.J. 451, 452 (1996).

23. UTAH CODE ANN. § 46-3-305(1) (Supp. 1997). The Utah Act does not define "reasonable care," and no guidelines or consensus exists over what steps are appropriate for protecting private keys.

24. These arguments are set out in detail in Comment, *supra* note 7, at 1167-86. *See also* U.C.C. § 3-403(a) ("an unauthorized signature is ineffective except as to the signature of the unauthorized signer").

Yet if the subscriber does not bear full liability under this scenario, where else would the loss fall? On the CA? They could neither prevent the harm, nor realistically insure against such indeterminate losses via pricing mechanisms. CAs presumably would not know whether a particular certificate was going to be used in a purchase of a piece of clip art or in a real estate closing. Thus, the CA could not charge a price that would be commensurate with the CAs corresponding risk of loss if the CA were to bear liability for fraud involving the certificate.

Could the loss fall on the relying party? The goals of a PKI would be undermined, and an opportunity for fraudulent collusion would be presented, if the relying party bears the risk.[25]

While the issues are perhaps less stark, this "liability trilemma" plays out similarly for each liability scenario present under open PKI. If CA liability is limited for erroneous certificates issued through no "fault" of the CA, who bears the risk of loss when fraud does occur? The relying party? The individual whose name is on the certificate, despite the fact that they may have no connection to the situation at all? What if CA liability is limited in the event of a compromised CA private key? Will the loss fall on any relying parties who are consequently defrauded? On subscribers who must revoke and replace certificates?

The fundamental assertion of this article is that there is no satisfactory solution to this problem under an open PKI model. Certainly there is no one-size-fits-all solution that can be imposed via legislation. As further described below, it is conceivable that market mechanisms may be able to solve this "liability trilemma" via contracts, and the "contractual privity problem" is no barrier to this result. However, this result is nonetheless unlikely: the open PKI model is an inherently flawed business model.

---

25. If relying parties bore the risk of loss, fraudulent collusion could occur when a subscriber willfully discloses their private key to a criminal knowing that, existing laws concerning fraud aside, the subscriber will not bear any resulting loss. Moreover, apart from intentional wrongdoing, subscribers would presumably have less incentive to be cautious with their private keys, more keys would be stolen, and, thus, the reliability of digitally signed documents would be increasingly suspect and the envisioned PKI would be underutilized.

V. THE "CONTRACTUAL PRIVITY" RED HERRING

The "contractual privity problem" as a justification for open PKI-oriented legislation is a red herring. At least one commercial CA, VeriSign, has emerged unsupported by legislation, and is largely pursuing the open PKI model. This CA is betting a substantial investment that they can form "webwrap" or "click-through" contracts with relying parties when the relying parties verify certificates. That is, when relying parties connect to the VeriSign Web site to determine whether a particular certificate has been revoked (placed on VeriSign's CRL), they are presented with a "click through" agreement which defines the limited warranties VeriSign offers to relying parties on certificates and which strictly limits VeriSign's potential liability. VeriSign enters into a similar agreement with subscribers when the subscriber first obtains a certificate.[26]

Click-through agreements are not without their problems. The question of whether they are enforceable at all is not definitively settled, and there can be other potential problems depending on specific circumstances.[27] However, this issue is not unique to the CA industry.

---

26. One aspect of VeriSign's approach is puzzling. VeriSign attempts to bind both subscribers and relying parties to legal terms detailed in its Certification Practices Statement (CPS). The *Digital Signature Guidelines* similarly envision a CA's CPS as setting out the basic legal relationship between CA, subscriber, and relying parties. DIGITAL SIGNATURE GUIDELINES, *supra* note 3, at 32-33. This approach is problematic. Certification Practice Statements are fundamentally misconceived as legal documents; they are full of technical information that simply is not legally relevant. Jumbling up CA-subscriber legal terms, CA-relying party legal terms, and complex technical information in a single document creates contract formation problems that are virtually insurmountable.

A more effective approach would be to use two short contracts: a CA-subscriber agreement, and a CA-relying party agreement. There would be only a few key provisions in each contract: warranties (both CA to subscriber/relying party as well as covenants from subscriber or relying party to CA), warranty disclaimers, liability limitations, indemnification, dispute resolution, and choice of law. These provisions would not have to be drafted in a particularly complex fashion. For example, there is no need for a CA to warrant that they will take steps A through Z, in numbing detail. Rather, a CA can simply warrant that a particular set of facts will be true, and then scale their liability for nonperformance of those promises through liability limitation provisions. A CA could incorporate a CPS by reference into their contracts, if they wanted to link their promises to a particular set of practices, but this really is not necessary. A CPS would better serve as a marketing document rather than a legal document.

27. *See* Carey R. Ramos and Joseph P. Verdon, *Shrinkwrap and Click-On Licenses After* ProCD v. Zeidenberg, THE COMPUTER LAWYER, Sept. 1996, at 1. *See also* Gary H. Moore and J. David Hadden, *On-Line Software Distribution: New Life For 'Shrinkwrap' Licenses?*, THE COMPUTER LAWYER, Apr. 1996, at 1 (noting, among other things that "the 'on-line' version of the ubiquitous 'shrinkwrap' license stands a far

Many online businesses are forced to rely on click-through agreements. Several recent court decisions have strongly suggested that they will be enforceable; a legislation-drafting effort underway is attempting to settle the question.[28] Click-through agreements present a mechanism by which CAs can attempt to allocate risk contractually.[29]

Accordingly, the open PKI model can compete in the marketplace in the absence of any special legislation. Independent commercial CAs can form enforceable contracts with both subscribers and relying parties, via click-through agreements, and, thus, allocate risk contractually. Various CAs utilizing this model presumably would compete on warranty, indemnification, liability limitation, and other contractual terms. This is not a winning business model, however.

Even utilizing the flexibility and dynamism of the market, CAs practicing the open PKI model will not be able to solve the "liability trilemma." In the long run, CAs will not be able to simultaneously offer certificates at reasonable prices, along with contract terms acceptable to both subscribers and relying parties. The open PKI model simply poses

---

greater chance of being enforced than its hard-copy cousin.").

28. Two recent federal appellate court decisions have upheld the enforceability of "shrinkwrap" agreements using a rationale which will also support the enforceability of "click-through" agreements. *See* ProCD, Inc. v. Ziedenberg, 86 F.3d 1447 (7th Cir. 1996); Hill v. Gateway 2000, Inc., No. 96-3294, 1997 U.S. App. LEXIS 176 (7th Cir. Jan. 6, 1997). The National Conference of Commissioners on Uniform State Laws and the American Law Institute are nearing completion of an amendment to the Uniform Commercial Code which will endorse click-through agreements. *See* (visited Jan. 17, 1998) <http://www.law.uh.edu/ucc2b/>; Carole A. Kunze, *A Guide to the Proposed Law on Software Transactions: Draft U.C.C. Article 2B- Licenses* (visited Jan. 17, 1998) <http://www.softwareindustry.org/issues/guide/index.html>.

29. A similar mechanism for allocating risk via contract mechanisms is discussed in David G. Masse and Andrew D. Fernandes, *Economic Modelling and Risk Management in Public Key Infrastructures* ¶ 178 (Version 3.0, Apr. 15, 1997) (visited Jan. 17, 1998) <http://www.chait-amyot.ca/docs/pki.html>. This thoughtful paper, presented at the RSA Data Security, Inc. annual symposium in San Francisco on January 31, 1997, touches on many of the same issues presented in this article, although it reaches some different conclusions. One potential hitch with the scenario described here is if a court honors a relying party's claim against a CA in a circumstance where the relying party failed to check a CRL, or where a relying party relied on an expired certificate which the CA no longer maintained on a CRL, and, thus, did not enter into a contract with the CA. This result seems unlikely, however, in light of the "unreasonableness" of not checking a CRL or relying on an expired certificate; a court would likely determine that such actions constituted an assumption of risk of loss by the relying party, at least as between the relying party and the CA. If this result were to consistently occur it could be dealt with in focused legislation.

too much inherent risk of loss that must be borne by one of these three parties.

VI. THE CLOSED PKI MODEL

To assert that open PKI is not viable is not to say that public key cryptography will not play an important role in electronic commerce. The closed PKI model offers some significant benefits, and will likely compete and win in the marketplace for use in a variety of applications.

A closed PKI has been defined as a system wherein a contract or a series of contracts identifies and defines the rights and responsibilities of all parties to a particular transaction.[30] This definition came about in response to the assumption that the "contractual privity problem" prevented contract formation with relying parties in an open system. As discussed above, this is an erroneous assumption: CAs can form contracts with relying parties in an open system. Thus, a better definition of a closed system is one where public key mechanisms are used within a specific, bounded context.[31]

Risk management is the critical area of difference between closed and open PKI. Within a bounded context the liability allocation problems which are intractable under the open model become manageable, primarily because potential liability exposure is quantifiable and limited in scope. For example, the proprietor of an online "mall" might issue certificates to potential customers and to merchants. The proprietor, acting as a CA, has the opportunity to enter into contractual relationships both with consumers and with the merchants who will rely on the certificates, and, thus, can allocate risk contractually. Moreover, the risk to be allocated is relatively small. Unlike under the open model, the CA knows exactly what the certificates that are issued will be used for. The CA can accurately predict and manage potential losses, and either absorb this cost via pricing mechanisms, or assign it to either subscribers or merchants by contract.

Similarly, a merchant might issue certificates directly to its customers. The owner of an online magazine, for example, might mail diskettes containing certificates directly to subscribers of the paper version of the same magazine. Such certificates could be installed on the subscriber's Web browser and used to access the online magazine, and perhaps to

---

30. This definition was used in the *Report of the ILPF Working Group on Certification Authority Practices*, at app. 2 (draft Feb. 24, 1997) (visited Jan. 5, 1998) <http://www.ilpf.org/work/ca/draft.htm>. For more information about the Internet Law and Policy Forum, see (visited Jan. 17, 1998) <http://www.ilpf.org>.

31. Note that some problems and alternative formulations exist with the open/closed PKI nomenclature. *See supra* note 20.

order related merchandise. The magazine vendor would be well positioned to determine whether such certificates would be sufficiently trustworthy for the purposes for which they were being used. Again, such a scenario does not implicate the difficult risk allocation questions associated with the open model.

Another example might be a business-to-business trading network. Businesses may have processes and equipment in place which enables them to carefully manage private encryption keys. They may thus be quite willing to agree to contract terms much more "onerous" than the terms imposed by the Utah Act, for example. A business might agree to be strictly liable for all documents signed with its private key, under the appropriate circumstances. A closed PKI system preserves this type of flexibility.

In addition to better risk management capabilities, the closed PKI model offers another benefit: the ability of a certifying entity to realistically certify *attributes* or *authority* as opposed to *identity*. As other commentators have described, the seemingly mundane concept of identity "is a very subtle notion whose nuances go to the very core of human social and economic interaction."[32] CAs practicing the open model are unable to capture these nuances, and instead bind public keys to identification information that is frequently irrelevant in an online transaction. *Who* a person is—in the sense of knowing, say, their name and address—may be less relevant than knowing whether they have authority to act on a corporation's behalf, or whether they are entitled to access a particular online resource, or whether an account they have is valid. The closed PKI model can manage both the logistics and risk associated with attribute or authority certificates; the open PKI model cannot.[33]

Stephen Kent gave a fascinating talk last year entitled "Let a Thousand (Ten-Thousand?) CAs Bloom,"[34] and his central theme deserves

---

32. Masse and Fernandes, *supra* note 29, at ¶ 106. *See also* Carl Ellison, *Establishing Identity Without Certification Authorities* (July 22-25, 1996) (visited Jan. 17, 1998) <http://www.clark.net/pub/cme/usenix.html>.

33. Masse and Fernandes, *supra* note 29, at ¶ 123; *see also id.* at ¶ 126 et seq. (discussion of "credential certificates").

34. Stephen Kent is a well-known and highly respected cryptographer affiliated with BBN Corporation. The presentation was given at the DIMACS Workshop on Trust Management in Networks on September 30, 1996. See (visited Jan. 17, 1998) <http://dimacs.rutgers.edu/Workshops/Management/program.html> for a description of

repeating: let anyone who wants to be a CA be a CA, certifying their own customers, employees, members, etc. Individuals do not need one overarching identity certificate, to be used in every conceivable circumstance. People can have many certificates, inconspicuously installed in a Web browser, each of which is used in a specific, narrow context. Within the confines of such a bounded context the risk allocation questions which are insurmountable under an open PKI model become much, much easier. Moreover, unlike the rigid, hierarchical structure of an open PKI, such a scenario embraces the chaotic, fast-paced environment of the Internet.

Market trends appear to support the conclusion that the closed model will be the winner in the marketplace. Recall the flurry of announcements concerning commercial CA services which took place in early 1996. VeriSign was spun off from RSA. GTE CyberTrust announced its imminent competing "CyberSign" service. Nortel began issuing demonstration sever certificates, and was rumored to be on the verge of launching full-blown CA activities. MCI, AT&T, and IBM all hinted at planned CA services.[35] The U.S. Postal Service announced its intentions to act as a CA.[36]

Look at the state of the commercial CA marketplace now. CyberTrust finally started issuing certificates in early 1997—but only for SET,[37] which presumably will leverage the risk allocation mechanisms already present in the bank card industry. The CyberSign service, based on the open PKI model, has evidently been abandoned in favor of CyberTrust's emphasis on their "customer branded service."[38] CertCo, a Bankers

---

the conference. Kent's presentation is summarized at (visited Jan. 17, 1998) <http://dimacs.rutgers.edu/Workshops/Management/Kent.html>.

35. For example, in early 1996 Netscape Navigator version 2.x came equipped to recognize certificates issued by MCI, AT&T, and others. Froomkin, *supra* note 4, at 112.

36. Shawn P. McCarthy, *Postal Service Can Time-Stamp Certified Mail As E-Mail, Safely*, GOVERNMENT COMPUTER NEWS, November 18, 1996, at 39.

37. SET is the Secure Electronic Transactions protocol promoted by Visa and Mastercard, among others. For links to basic information about SET, as well as a thoughtful and provocative criticism of the SET effort, see Simson Garfinkel, *Is the Web Set for SET?*, HOTWIRED: PACKET (June 18, 1997) <http://www.hotwired.com/packet/garfinkel/>.

38. In response to an early draft on this article, Tom Carty, Vice President of Business Development & Marketing for GTE CyberTrust, e-mailed the author and wrote:

> GTE strongly endorses the closed PKI model. This has been reflected in CyberTrust's architecture from the beginning in the form of Customer Branded Service which is based on or Virtual Certification Authority (VCA) technology. This closed form of PKI service has been adopted by a number of our customers and is well suited to many of the business needs today. We believe business customers have established relationships with many of their customers and are in the best position to determine who should receive a certificate for

Trust spin-off and a newer entrant into the CA industry, appears to be focused largely on a contract-oriented model.[39] The rumors of Nortel's (now Entrust[40]) impending entry into the CA business proved false. Entrust's services focus exclusively on back-end support for other companies that want to offer CA services. Likewise, IBM's offerings focus on back end support.[41] The Postal Service's plans never materialized. All of these developments are consistent with the dominance of the closed PKI model. Industry players are focusing their attention on "letting ten-thousand CAs bloom."

VeriSign is the only North American company that is actively pursuing the classic open PKI model as a business model. Indeed, judging from their press releases they are doing so with some success, and they are planning innovative ways of addressing problematic aspects of their business model, such as via a recently-announced limited insurance plan for subscribers. However, even VeriSign used the high-profile forum provided by the 1997 RSA Data Security Conference to announce its "Private Label Digital ID Services," which are back-end support for companies that want to offer CA services.[42] Thus, even VeriSign is focusing on offering its services to other companies on an outsourcing basis, consistent with the closed PKI model.

VII. OPEN PKI LEGISLATION LIVES ON

Despite increased recognition of the problems associated with the open PKI model, legislation enshrining this model in law and regulation continues to be proposed and enacted. This can partly be attributed to factors that could be labeled "psychological." The rigid, straightforward hierarchies inherent in the open model are likely deeply appealing to the

---

operating with them as business clients.
E-mail message from Tom Carty, Vice President, GTE CyberTrust, to C. Bradford Biddle, (May 26, 1997) (on file with author). Other senior executives from companies active in the certification authority industry also responded with similar comments.

39. *See* Chris Jones, *Banking Spin-Off to Issue Digital Ids: CertCo's Software Will Distribute Electronic Transaction Risk*, INFOWORLD, Jan. 13, 1997, at 21.

40. *See* (visited Jan. 17, 1998) <http://www.entrust.com> for more information.

41. *See* (visited Jan. 17, 1998) <http://www.internet.ibm.com/commercepoint/registry/index.html>.

42. *VeriSign Provides Custom Digital ID Services to Large Corporate Customers: NOVUS Services and Toppan Printing of Japan Among Those to Select VeriSign to Provide Digital Authentication Services for Internet Customers*, PR NEWSWIRE (Jan. 27, 1997) <http://www.verisign.com/pr/pr_large.html>.

sensibilities of lawyers and legislators as they contemplate the often chaotic and inscrutable Internet. Moreover, enacting legislation—even legislation that is not well-understood—satisfies the legislative urge to "do something" in the face of a rapidly changing economic environment.

Additionally, legislators may be under the mistaken impression that special legal rules are needed to accommodate electronic commerce. In fact, however, the law is quite flexible and supportive of new commercial methods.[43] The oft-heard assertion that enforceable electronic contracts cannot be formed in the absence of a legislatively-endorsed digital signature is simply wrong.[44] As a general matter, legislation is not needed in order to accommodate public key cryptography or other emerging authentication technologies. The few areas where existing legal rules impede electronic commerce can be addressed with narrow, targeted legislation.[45]

Another factor behind the continued momentum of open PKI digital signature legislation is less benign: the infrastructure created by such legislation is ideally suited for implementation of a key escrow scheme. Indeed, such an idea was initially broached by the Clinton Administration in a report released in May of 1996. This report described a vision for a PKI consistent with the Utah/ABA Guidelines model, and noted:

> To participate in the network a user needs a public key certificate signed by a CA which "binds" the user's identity to their public key. One condition of obtaining a certificate is that sufficient information (e.g., private keys or other information as appropriate) has been escrowed with a certified escrow authority to allow access to a user's data or communications. (As noted before, this might be the CA or an independent escrow authority) . . . .[46]

---

43. *See* Benjamin Wright, *Electronic Commerce Legislation: Frequently Asked Questions*, CYBERSPACE LAWYER, Apr. 1997, at 10. *See also* Peter N. Weiss, *Security Requirements and Evidentiary Issues in the Interchange of Electronic Documents: Steps Towards Developing a Security Policy, reprinted in* XII JOHN MARSHALL JOURNAL OF COMPUTER AND INFORMATION LAW 425 (1993).

44. This point is made quite effectively, and examples of the misinformed conventional wisdom are given, in BENJAMIN WRIGHT, THE LAW OF ELECTRONIC COMMERCE, EDI, E-MAIL, AND INTERNET: TECHNOLOGY, PROOF, AND LIABILITY §§ ET 1.1, 1.4, 1.5 (2d ed. 1995 & Supp. 1996).

45. The draft legislation in Massachusetts provides an example of legislation narrowly tailored to provide clarity concerning issues such as signature, writing, and records requirements. *See Massachusetts Electronic Records and Signature Act* (draft Nov. 4, 1997), (visited Jan. 17, 1998) <http://www.magnet.state.ma.us/itd/legal/mersa.htm>.

46. Executive Office of the President, Office of Management and Budget, *Enabling Privacy, Commerce, Security and Public Safety in the Global Information Infrastructure* (May 20, 1996), (visited Jan. 17, 1998) <http://www.epic.org/crypto/key_escrow/white_paper.html>.

Draft legislation designed to implement such a plan was released in April, and a bill was introduced in the Senate in June.[47] Similar legislation was introduced in the United Kingdom.[48] The profound civil liberties concerns implicated by such legislation, and the negative effects on commerce of the policies underlying key escrow, are well documented.[49]

VIII. CONCLUSION

The open PKI model envisioned by many existing and proposed digital signature laws is not viable. This legislation presumes a business model that cannot internalize the costs associated with its implementation. In attempting to solve an unsolvable problem, current digital signature laws shift an immense liability burden onto consumers who use the infrastructure envisioned by these laws. Consumers would reject this result in any true marketplace transaction.

Digital signatures will play a significant role in electronic commerce, but under a closed PKI system where the liability problems associated with digital signatures become manageable. The open PKI model can and should compete against closed PKI and other authentication technologies without the benefit of special legislation. It will not, however, be the winning business model.

Digital signature laws which impose a particular view of electronic commerce on the marketplace should be abandoned. The time for legislation and regulation is after identifiable problems exist in a mature industry, not before an industry even exists. The existing legal infrastructure can accommodate new technologies without dramatic new legislation. Premature legislation and regulation risks creating market distortions which can prevent the market from arriving at much better results than those envisioned by governmental policymakers. Certainly this is true in the case of open PKI: digital signature legislation based

---

47. S. 909, the McCain-Kerrey "Secure Public Networks Act," was approved by the Senate Commerce Committee on June 20, 1997. Secure Public Networks Act, S.909, 105th Cong. (1997) (enacted). *See also* (visited Jan. 17, 1998) <http://www.cdt.org/crypto/> (containing details concerning the draft legislation released in April and for updated information concerning S. 909).

48. *See* (visited Mar. 26, 1997) <http://dtiinfo.dti.gov.uk/pubs/>.

49. *See* A. Michael Froomkin, *It Came From Planet Clipper: The Battle Over Cryptographic Key "Escrow,"* 1996 U. CHI. LEGAL F. 15, and sources cited therein.

on the Utah/ABA Guidelines model imposes a business model which could not survive under the discipline of the marketplace.