Digital Signature as THE Electronic Signature

Abstract

Worldwide, countries and sub-jurisdictions are enacting legislation that gives legal status to electronic signatures. The primary motivations for doing so are 1) to encourage the development of electronic commerce by removing the perceived barrier presented by its nebulous status and 2) to establish the 'trust' and predictability needed by parties doing business online. Although this is a first big step, legal status alone is not sufficient. In this paper the discussion will begin with the functions of a signature, the various definitions of electronic signatures, and the ensuing trust elements that are essential to their usefulness in an e-commerce[1] setting. The case is then developed that in order to meet all of the pre-conditions for validity and trust that business practices and legislation require, a digital, public key-based, signature is not only a strong candidate but the only comprehensive and practical electronic signature mechanism available today.

1 Introduction

The driving need to replicate an inked signature in the world of electronic commerce is an obvious one. We cannot transact business nor risk enterprise dollars and product without assurances that the agreements entered into will be honored in the same fashion they are in the purely brick and mortar world. The need for electronically signed documentation is ubiquitous;

- Electronic business contracts must be verifiable (source and content);
- Evidence in court cases, in the form of digital data, must have integrity that can be verified;
- Access to sensitive records must be controlled and auditable; the chain of trust for archives of signed data must be maintainable;
- Sessions for online purchasing, banking, and access to services require authentication;

[1] For the purposes of this discussion, e-business refers to any business conducted using the efficiencies of the Internet. It includes, but is not limited to, government, finance, healthcare, and all commercial vertical market applications.

- Content protection in general, i.e. how to make information available through public channels while maintaining control over it;
- Software itself, the engine behind the digital explosion, is an example of content that requires trusted processes to assure it authenticity and integrity to a relying party.

In the context of this paper, the terminology “electronic signature” will be the generic technology neutral term and “digital signature” will be reserved for the specific mechanism that is based on public key cryptographic technology. Confusion can arise in reading legislation because this distinction is not always maintained. However, within the security industry and in the context of e-business enablement the distinction to be applied here is the commonly accepted one.

Electronic signatures cannot solve all the problems associated with an inked signatures but it is reasonable to expect that they meet the challenge of solving problems the electronic medium introduces. In fact, when digital signatures are the technology employed, they increase the trust in the overall concept of a signature as well.

As will be discussed below, much of the legislation regarding electronic signatures differentiates between what constitutes an electronic signature and what constitutes a legally valid electronic signature. Even when the pre-conditions of validity are stated, *the means to achieve them* are not well defined

2 The Definition of a Signature

One generally accepted definition of a signature as written in the Uniform Commercial Code of the United States is “any symbol executed or adopted by a party with present intention to authenticate a writing.” That “intention to authenticate” encompasses a variety of actions all relative to the context of a given document:

- Acknowledging that the signer has had the right to read the content,
- Affirming the content has been read, or
- Confirming obligation to, or authorization of, the content.

The UCC definition makes no reference to ink or paper and in fact case law in the U.S. supports that it is the signatory’s *intent within a context* that is critical to this notion as defined.

However, without taking steps to extend this definition explicitly, legislators perceived the existing definition to be an impediment to the advancement of e-commerce because of the thousands of statutes and regulation that require transactions to be signed and seemingly in writing. As will be seen in some of the definitions of an electronic signature we review, the legislative drafters were also concerned about the risks introduced by the electronic nature of the data. Consequently, much of the legislation places greater burdens on the electronic version of a signature than were ever imposed on its inked counterpart although the function of the signature remains the same and is focused on the intent of the signer.

3 Legislative Response

Globally the legislative response has been huge albeit inconsistent in its approach. The following list is not inclusive and only meant to highlight the level of activity.

- U.S. Government
  - Federal Electronic Signatures in Global and National Commerce Act of 2000
- 15 countries including Canada, Singapore, Japan and Australia [LEG]
- European Union Directive on a Common Framework for Electronic Signatures- March 1999
- Forty-nine U.S. States [LEG]
- National and international organizations including
  - The National Conference of Commissioners on Uniform State Laws (NCCUSL)Uniform Electronic Transaction Act – UETA 1999United Nations Commission on International Trade Law (UNCITRAL)Model Law on Electronic Signatures 2000Organization for Economic Cooperation and Development (OECD)Guidelines for Cryptography Policy 1997

The rationale for most of the legislation is based on one or both of the following:

- *Enabling* legislation to promote the development of e-commerce by removing a perceived barrier; and to highlight the need for trusted process and predictability;

- *Regulatory* legislation to establish preconditions for the validity and enforceability of an electronic signature over and above those imposed on an inked signature.

Although the judiciary will certainly determine the rules that govern on-line business transactions, early legislation was designed to keep the e-commerce ball rolling and to encourage technologists to find solutions to these critical legislative and business requirements.

The inconsistency in the statutes can undermine the trust and predictability they set out to provide. What qualifies as an electronic signature, when is it legally binding, what makes it trustworthy, and what transactions can be undertaken (or are excepted) are addressed separately and differently. Some legislation attempts to create predictability in the process, and some imposes legislative structures on particular solutions, but for the most part, they remain silent on many of these important questions. Some statues are binding while others are meant to fill the void in the absence of contractual arrangements between the parties.

What follows here is an attempt to highlight the major characteristics of existing legislation. There are a number of websites that are excellent sources of up-to-date information on the specifics of global legislative activity for the interested reader [LEG] Our focus here is not so much on the specifics of any single law as on the commonalities and thinking behind the legislation in general.

3.1 Three approaches

Legislation has taken three basically different approaches in defining what qualifies an electronic signature to meet legal and regulatory requirements for signed documentation, when any legislation exits at all. [SME] These approaches are a function of the effect the legislation was to have.

3.1.1 All qualify

Most U.S. states as well as the Federal Electronic Signatures in Global and National Commerce Act of 2000 have taken the position that, like its inked counterpart, any electronic signature can qualify as a legal signature. The definition of an electronic signature in those states are all similar to Arizona's in scope: *"Electronic Signature" means an electronic sound, symbol or process that is attached to or logically associated with a record and that is executed or adopted by an individual with the intent to sign the record*. This definition is essentially parallel to that of an inked signature in the Uniform Commercial Code, as described above. There is no requirement that it even identify the signer. This legislation is enabling in that it explicitly states what was already implicitly true.

In the legislative responses from Illinois, South Carolina, and Singapore, while still recognizing the enforceability of *any* electronic signature that meets the definition, the concept of a 'secure' signature is developed and given special legal benefits. The purpose of this is to encourage the development of signature mechanisms that consider the importance of trust in the process. It also allows contractual agreements to prevail while encouraging some predictability in the absence thereof.

Using Illinois as an example, a secure electronic signature is one that has:

- A characterization agreed by the parties, **or**
- Technology certified by the Secretary of State to produce an electronic signature that:

- Is unique to the signer *within context;*
- Objectively identifies the signer;
- Is reliably created by such identified person;
- Is linked to the data record such that any intentional or unintentional change to the data or the signature invalidates the signature.

The legal benefit bestowed on these signatures is a rebuttable evidentiary presumption that the signature belongs to the person with whom it is associated. This is still only enabling legislation because the added conditions are not required for enforceability. The legal benefit is the carrot to encourage the use of more trustworthy processes.

3.1.2 Security Attributes Required

A third of the U.S. states as well as the European Directive augment the basic definition with 4 security attributes that are preconditions for legal enforceability and trustworthiness. An electronic signature must have controls that assure it is:

- Unique to the person using it;
  - Is this absolute uniqueness or uniqueness within a context?
  - What third party assures the binding of the person to the signature and makes it unique?
- Capable of verification;
  - Current and/or persistent validation?
  - How is a signature's full electronic path defined?
- Under the sole control of the person using it;
  - Does sole control apply at all times or only at the time of use?
  - Does sole control refer to the signing 'device' and if so how is that determined?
  - What about "authorized" use by another?
- Linked to the data in such a manner that if the data is changed, the signature is invalidated.
  - This implies a cryptographic process that uses the bits of the data in creating the signature but rarely is this mentioned. The Japanese definition of an electronic signature requires that it use a cryptographic process.

These preconditions for validity are applied to electronic signatures only, not to inked signatures. As stated, they do make sense in the electronic environment. Unfortunately the meaning of these seemingly straightforward requirements is vague and how to achieve them is not explicitly defined in the regulations. Comment is rarely even made as to whether by contract or agreement, the parties involved may dispense with these conditions.

The UNCITRAL Model Law imposes a slightly different set of basic requirements:

- Identity of the signer must be included;
- The signer's approval of the content must be indicated;
- The method must be as reliable as was appropriate for the purposes for which the message was generated or communicated.

These too are simply stated conditions on what constitutes a legally binding signature. Similarly here no guidance is given as to which technologies might meet these requirements and what criteria will be used to measure compliance.

3.1.3 Digital Signatures

Five US states - Minnesota, Missouri, New Hampshire, Utah, and Washington - have chosen not to maintain a technology neutral approach and have designated digital signatures as the only authorized electronic signature. In Hong Kong and in Canada, although a very broad definition for an electronic signature is given, only a digital signature based on asymmetric cryptographic principles satisfies the requirements of law. In Singapore, although a digital signature is not required, it is clearly being promoted as preferable. There are standards and controls established for licensed Certificate Authorities (CA) and limitations on liability granted for those CAs that comply.

3.2 Legislative Forcing Function

Although most statutes do not explicitly address the issue of trust, trust is in fact the fundamental “forcing function” for these various legislations. In order to encourage e-commerce, a legally binding electronic signature must be defined. The statutes that follow the approach discussed in 3.1.2 above make legality contingent upon satisfying the business needs for trusted processes. Even when the definition is left broad, the rules of law will impose these same requirements on any electronic signature that is to have value in an evidentiary proceeding. Meeting the requirements for trustworthiness thereby confers status as an enforceable signature for the purposes of the statute’s signature requirements. These attributes are considered essential for trust in the business process and stand as benchmarks to be met. How rigorously or broadly they should be construed, however, is left open.

As we have seen, other jurisdictions allow that *any* electronic signature (meeting the basic definition) *can* meet legal requirements, but simultaneously recognize that some electronic signatures are more trustworthy than others. This thinking mirrors the paper world where a handwritten signature is more trustworthy than an “X”, and a notarized signature is more trustworthy than either. Again the criteria for trustworthiness vary between statutes as does the situations under which they apply. Those that qualify are afforded a rebuttable evidentiary presumption that the signature belongs to the person with whom it is associated. The purpose of outlining these qualifications within the statutes is to encourage the use of more secure mechanisms. This is a commendable objective but still leaves more questions than answers.

For those legislative bodies whose statutes focus on digital signature technology as the sole technology afforded this presumption, the statutes impose regulatory requirements on the certificate authority and require them to be licensed. Any messages verifiable using certificates issued by these ‘trusted’ authorities, are then considered trustworthy.

Although the legislative responses are inconsistent, where there is more than summary thought given to the signature requirements, one thread is common and clear: the recognition that for e-commerce to proceed, the authenticity, integrity, and non-repudiation of transactions must be achieved in a trustworthy (secure) fashion that gains the confidence of all parties involved.

4 Trust Components of Electronic Signatures

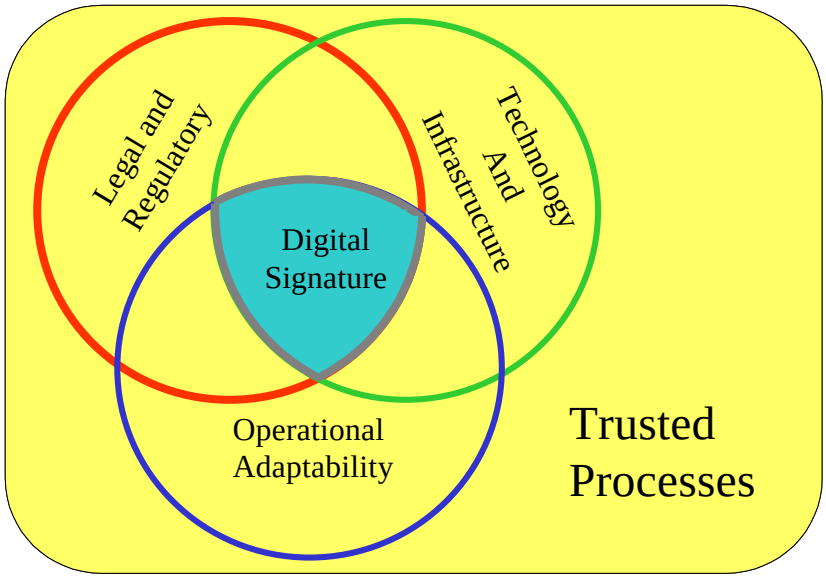

Figure 1: Digital Signatures are at the intersection of the Trust requirements.

It is said that a picture is worth a 1000 words. The solution space for electronic signatures includes trusted processes that satisfy external legal requirements as well as internal business policy requirements, and utilizes technologies that are well vetted and up to the job. Since much of the legislation has chosen to remain technology neutral up until this point in time, it is useful to discuss the trust issues that underlay the requirements as set forth and use them to discuss the attributes of various technologies to determine whether they might meet those requirements. Some of these trust issues are the same as those associated with an inked signature and some are introduced by the electronic medium itself. It is interesting to note however that when the legislative response is more than minimal, the trust burden on the electronic signature far exceeds that placed on its handwritten counterpart.

4.1 Level of Trust Required

The need for a trustworthy signature can be used as one measure of risk tolerance. How much loss rides on the contract or transaction? If the dollar value is high or the potential for property or reputation loss is high, the ability to achieve non-repudiation becomes essential. “Better, faster, cheaper” is the business motto and secure transactions cost more and can be slower and demand more processing power. How much risk is there in the transaction? As the risk grows so does the need for trusted processes. At one end of the spectrum might be password identification to a website. If the purpose of the password is loyalty and the benefits of access would not encourage abuse (easier to get your own), then this becomes a very acceptable ‘signature’ mechanism. Another example is the simple mouse click that a user of software makes to accept the licensing agreement before the software will install or operate. This may be a sufficient signature given that, if the mechanism fails safe, there is no transfer of information or property unless there is acceptance. The individual identity is of less importance. The presumption of acceptance remains as long as the software is installed and operating.

Financial institutions and credit card companies are experts at assessing risk and determining the cost of fraud. Where the potential loss associated with individual transactions is comparatively small as it is in retail sales, the need for a non-refutable signature may be of less importance to the institution. But even here, as the number of fraudulent transactions becomes large due to the nature of the Internet and the ability to attack broadly and quickly, building the cost of fraud into the cost of doing business will reach its limits. The assurance must be built into the process *before* the transaction takes place.

For the remainder of the discussion, we assume that this need is present, that the ecommerce transactions are of a volume or value that a trusted process is essential in order for the efficiencies of the Internet to be realized in the conduction of business transactions.

Digital Signatures Perform

For the bulk of commercial, banking, and government transactions, strong authentication of participation in the execution of a transaction or contract is imperative. As the required level of trust increases, the complexity of the services needed to assure the trust increases as well. The discussion extends far beyond just the signature mechanism at hand. For the business process to be sustainable, the infrastructure that manages the signature is what provides the trust. We begin to see that the legislation is really addressing this issue as well. Digital signatures have a well vetted technological infrastructure behind them that is based on digital certificates that allow us to deal with issues such as efficient but secure distribution of the cryptographic keys, binding the identity of the owner to his key, real time verification of the signature by the relying party, and near real time verification of the identity and validity of the ownership of the signature. Since the mid-70's, the infrastructure has been in the process of refinement. The concept of a signing mechanism was seen early on to be viable only in the context of well-managed security services

4.2 Authenticity, Integrity, and Non-repudiation

After determining whether or not an electronic signature can be considered legal, the next fundamental question is: does it help establish trust in the content and origin of the transaction, data, or contract of interest. In the brick and mortar world, personal relationships, face-to-face contract signings, notaries, and third party counsel are used to help establish trust in this most important aspect of conducting our business. As the reliance on paper shifts to electronic transactions and documents, so must the reliance on traditional trust factors shift to electronic security measures to ***authenticate*** the identity of our electronic business partners, customers, and suppliers before engaging in the exchange of information, goods, and services. In the event that an electronically signed contract is challenged, the need to establish the signatories to the contract is as real as with a conventional paper contract. This is the trust issue that gives rise to requirements for "uniqueness" and "sole control of the signature" as discussed earlier.

Another trust issue concerns establishing the ***integrity*** of content. Licensing digital content, online contract formation, financial transactions, and data archival all depend on the parties' trust that the content is accurate, in agreement with the terms known to the parties, and unalterable at some future date without the parties being aware of the change. Not only is this an issue for dispute resolution and evidentiary proceedings, but also for the good faith necessary in the everyday performance of business. Thus we see the requirement on a trustworthy electronic signature that the signature be "linked to the data in such a manner that if the data is changed, the signature is invalidated".

Non-repudiation is defined as authenticity and integrity provable to a third party, which implies that the signer cannot deny, at some later date, that they were a signatory to the content of an agreement, transaction, or contract. Non-repudiation is a function of an authenticated signature and unaltered content within a context. In the paper world, these assurances come from many sources together: paper with watermarks and special indicia that underlie the content, notarized or witnessed 'in person' signatures, sealed envelopes delivered by trusted third parties, and personal acquaintance. Even with all of these mechanisms available, disputes arise and non-repudiation can be elusive. In cyberspace, the parties frequently have no proximity to each other, making the reliance on computerized security measures such as trustworthy electronic signatures very appealing. Non-repudiation is still not guaranteed but with the correct signature scheme in place, the risk can be manageable.

Digital Signatures Perform

Passwords, biometrics, and private keys can all be argued to be under the sole control of the user. Biometrics and private keys can both claim to be unique and be verifiable. However, the requirement that the signature be "linked to the data in such a manner that if the data is changed, the signature is invalidated", requires some sort of cryptographic process to bind the two pieces of information electronically. The level of complexity jumps immediately with this last requirement and yet contractual processes, financial transactions, and legal non-repudiation require this leap as business processes are transformed to adapt to cyberspace. Utilizing cryptographic techniques necessitates the introduction of keying mechanisms. The binding of the signature to the data can be accomplished using either a (symmetric) shared secret key methodology or using digital signatures that exploit private/public key pair technology. Secret keys are not conducive in applications where verification is an important capability. If the key that is secret to the signing process must be distributed to everyone who needs to verify the signature, the uniqueness and sole control requirements are not achievable. With key pairs, only the signer knows the private signing key and all who wish to verify the signature may do so with the freely distributed public component of the pair. Authentication and data integrity are achieved, non-repudiation is enabled, and the pre-conditions of legality are met.

4.3 Verifying identity and trust in the verifier

Verification of the signature and its owner is of even more importance. Because the parties to many transactions are separated by distance and are personally unknown to each other, the questions begged are

- What party will serve to do the 'digital introduction'?
- What are the credentials of this third party?
- How do I use them to verify the signature?
- Is the process well vetted and backed by sound policy?
- How well does the third party 'know' the owner of the signature?
- Do the schemes of verification, of both the identity of the signature owner, and the signature itself have high assurance policy and procedure behind them?

These are some of the questions underlying the legislative requirement for the trustworthy signature to "be easily verified". This further points out the importance of the infrastructure. Any signature method we choose is only as good as the trust in the system that supports it. Reliable identity is critical.

Digital Signatures Perform

More than just an introduction, what is needed from a trusted third party is high assurance that the person presenting a given signature has the right to present it (i.e. is the owner of the certificate). Many of the current electronic signature statutes that add minimal additional requirements (e.g. UNCITRAL) require that the identity of the signer be clearly associated with the signature. The science behind the use and operation of a Certificate Authority (CA) has this task as its primary objective. It is a CA's responsibility to tie the identity of the entity (person, organization, machine) to the credential used by it to sign digital correspondence. The Certicate Practice Statement should clearly delineate the methodology used to establish the linkage before the certificate is issued. Further it is the CA's responsibility to notify relying parties if the binding is no longer valid (e.g. in the event of loss or compromise) and to establish under what circumstances the credential can be validly used. Public Key Infrastructures have a variety of trust models adaptable to the modern business models that drive their use. The operator of the CA adds to the trust of the process. Governments, vertical market leaders, and commercial entities that establish their reputation for integrity and scrupulous independence are strong candidates for applications that extend beyond a single enterprise. This linkage is as critical to the electronic signature as it is to the written signature. Without a strong trusted solution to this identity assurance issue, legal non-repudiation isn't possible.

4.4 Liability/responsibilities of the parties

Nothing comes with guarantees but in the paper world most business process relies heavily on time-honored rules of engagement and behavior. The likely consequences of particular activities are known because the rules are established through years of encounter and the consequences have a body of law behind their analyses. Cyberspace introduces risks that have not been encountered before. The crimes will remain the same but how they are carried out and the size of the threat they portend are changing dramatically. Rules of conduct in the face of these crimes are not yet well established. Efforts to clarify liability further add to the sense of trusted processes and enhance the perception of predictability.

Digital Signatures Perform

The roles and responsibilities of the stakeholders in the operation of a PKI are model dependent and have been a topic under development in government, banking, and legal communities for years now. One of the critical components of a system to manage trustworthy digital signatures is the body of policies that lay out how the CA and the certificates will be managed. Certification Practice statements and Certificate Policies detail the legal use of the signature, the liabilities accepted or denied by the CA, and the responsibilities of the subscriber and relying parties. Guidelines based on different models have been developed by the ABA [ABA] and NACHA [NAC]. Although these responsibilities will be dependent on the business drivers at hand, there is a voluminous body of knowledge available from which to draw. These guidelines and policy requirements form a solid basis for contractual arrangements to provide predictability.

4.5 Operational Feasibility

Certainly scalability and adaptability to the operational environment must be considered when deciding on a solution to the electronic signature dilemma. Because multiple applications demand the same services, the signature solution deployed should also be flexible enough to adapt to a variety of uses and remain trustworthy as it scales.

Digital Signatures Perform

The infrastructure that provides the trust for a digital signature is readily expanded to meet the variety of business applications that require its services. Any signature solution will incur institutional overhead and direct costs to the relying parties and even to the subscribers. The efficiencies they enable, however, are business critical, the benefits far outweighing those costs. See the PKI Forum paper *PKI and Financial Return on Investment* [ROI].

For that expenditure, the benefits include:

- The use of open systems to add efficiency and cost savings to modern global business processes
- Shorter business cycles which improve revenue streams
- Reduction in the risks posed by imposters
- Reduction in fraud, forgery, and false claims of altered transactions
- Compliance with legal requirements
- The only key management solution practical in a large scale transient (online) environment.

Cryptographic mechanisms are the only technical solution that can provide authenticity, integrity and non-repudiation for any type of data in a transient environment. Digital signatures are an extension of a PKI and carry with them the benefits of the trusted processes of a well-vetted technology that scales.

5 Conclusion

Signatures are a fact of life in order fulfillment, business commitments, contract negotiation and closure, financial agreements, and government compliance. In cyberspace, they are also imperatives for access to services and systems, for meaningful archival of digital documents where persistent signatures are a necessity, and for identifying a person before privileges can be invoked. Before e-commerce can take full advantage of the efficiencies and cost benefits of electronic transactions, trust in those transactions and in the commitments they imply must be assured. The regulatory endeavors to foster this trust are in early stages but developing. They have given the electronic signature legal status and have provided security attributes that provide a guide to the qualities that are fundamental to the trust requirements that business processes seek.

Digital signatures and the trust provided by their underlying infrastructure are based on a well vetted technology. Since the mid 80's the capabilities inherent to this discussion have been under analysis. Working models and standards have been established to support the development of a system, based on cryptographic principles, to provide secure authenticated data transfers that can be validated even when persistence is required. The issues of uniqueness, sole control, verification, identity management, and non-repudiation are the ones that the digital signature infrastructure set out to address. Other cryptographic solutions are possible but they are cumbersome in large scale transient environments and do not benefit from years of affirmation. The current literature on legislative trends shows that there is recognition of the value of this comprehensive technology even when its implementation is not a hard requirement. There is no other single technology to be named that meets both the preconditions of legal validity for a trustworthy signature and the business requirements for scalability, flexibility, and trust. And so the conclusions are compelling. The growing awareness of the need for ubiquitous authentication in electronic business can only highlight the benefits of digital signatures. Web services, federated identity management, XML and SAML based services all require a trusted scalable infrastructure in which to operate securely. Realistically no other solution today has the technical capability to deliver on this.

The inconsistencies in the statutes can undermine the trust and predictability we seek. It is imperative, as we move forward, to give more thought to the reasons behind the regulations so that coalescence to more precisely defined requirements can be achieved. Taking a technology neutral stance cannot be used to avoid addressing critical issues presented by new technologies nor to prejudice against the use of technologies that lend themselves so well to secure e-commerce.

Digital Signatures Perform

6 References

[ABA] http://www.abanet.org/scitech/ec/isc/home.html

[LEG] link to OASIS/PKI Forum links page for legal issues, currently
http://www.pkiforum.org/resources.html#pkilegal

[NAC] NACHA Electronic Payments Association http://www.nacha.org

[ROI] link to the OASIS PKI paper on ROI, currently
http://www.pkiforum.org/pdfs/Financial_Return_on_Investment.pdf

[SME] http://profs.lp.findlaw.com/signatures/signature_2.html