4 Palmers Green
St Johns
Worcester

1st April 1999

Stephen de Souza CIID/DTI
Room 220, 151 Buckingham Palace Road
London SW1W 9SS
(e-mail: sec@ciid.dti.gov.uk)

Dear Mr de Souza,

BUILDING CONFIDENCE IN ELECTRONIC COMMERCE

My comments on the government's consultation document on electronic commerce are set out in this letter and its attachment. I welcome the government's stated objective of promoting electronic commerce but I remain uncertain that its proposals will achieve this aim. This letter sets out my main concerns while the attachment covers more detailed issues, mostly those of a technical nature.

Non-Repudiation

In the UK, a signature is not binding if the alleged owner did not make it. The relying party hence has to show that a disputed signature is valid. The government's proposals seek to reverse this situation for digital signatures by introducing a rebuttable signature – one for which the signature owner has to show that a disputed signature is false. The technology to effectively support such a shift in the burden of proof is not available and this means that those seeking to use digital signatures may carry risks that have previously been carried by others (for example, by banks). This will mean that a digital signature certified by a licensed authority may actually carry more risk than one obtained where such provisions do not apply. This proposed change seems much more likely to undermine confidence in electronic commerce rather than to promote its development.

It would hence be preferable to promote electronic commerce using the approach proposed for Australian legislation where it is suggested that:

> "Unless otherwise agreed between the purported sender and the recipient of an electronic communication, the purported sender of the electronic communication is bound by that communication only if the communication was sent by the purported sender or with the authority of the purported sender."

Such an approach would ensure that provisions for electronic signatures match those for hand-written ones.

Liability

Any proposal to deal with liability issues for digital certificates in a way that is divorced from applications is unlikely to be effective. Digital certificates will be used in widely differing circumstances and the liability issues that arise will be very different. For example, while a false certificate might only jeopardise a low value financial transaction, such a certificate for a doctor's signature might put a patient's life at risk. The liability issues here are very different in both character and scale.

In a somewhat different context, if an organisation such as, for example, the Association of British Travel Agents (ABTA) were to use digital certificates to underwrite its member's digital signatures, this organisation should be free to set the terms of the promise that such certificates provide for consumers.

Moreover, in many applications the value of certificates will not be determined by liability when things go wrong but rather by the high percentage of certificates that prove to be correct.

These factors suggest that it is unrealistic to expect to cover liability issues for digital signature certificates in a generic way divorced from the context in which they will be used.

Consumer Risks

The financial risks that consumers face in electronic commerce would be most easily tackled by extending the protection afforded to credit card purchases to all electronic transactions. There would be a cost involved but this could be controlled in the way that credit card risks are now managed through effective vetting by banks of those who they allow to trade electronically under their terms and conditions.

There are significant privacy concerns in electronic commerce but these are almost exclusively ‘second party’ concerns related to the protection that traders provide in the handling of consumer supplied information. To promote electronic commerce the government hence needs to set out the data protection responsibilities of those who trade electronically whilst ensuring that the Data Protection Registrar has the legal powers, the resources and the technical expertise needed to make such provisions effective.

In comparison with these concerns, third party confidentiality services will carry very little interest for most consumers. In practice, it is hard to see a significant market for such services since third parties are not necessary for confidentiality provision. In consequence, it makes little sense for the direct parties to a shared secret to increase the risk of compromise by unnecessarily sharing it with others.

Law Enforcement

Much government thinking in terms of Law Enforcement requirements is based on trying to prevent or exercise control over the use of encryption for confidentiality purposes. Such activities are not only futile but potentially very damaging since they encourage the police and others to place their faith in solutions that will not work instead of investing in achievable goals.

In reality, the widespread use of encryption is inevitable whatever the government does and this means that law enforcement authorities have to develop the expertise required for their continued effectiveness in this new environment. In short, the police need to be better than criminals are when operating in cyberspace – there is no other solution and to pretend otherwise is dangerous because it undermines investments towards this end. In particular the government should now consider the establishment of a National Computer Forensics Laboratory so that law enforcement authorities have access to the best computing expertise available. With appropriate provisions for accountability and scrutiny, there is every reason to believe that both the police and the public would support such a development.

In view of the government announcement that it is reviewing such matters, it would be sensible to remove law enforcement requirements from these electronic commerce proposals so that they can be properly considered as a part of this review process.

Cryptography Export Controls

Electronic Commerce is international in scope and this means that it cannot succeed while an enabling mechanism – cryptography – is subject to export controls. In fact, export controls on civil cryptographic products are not consistent with the international arrangement under which they are supposedly justified (the Wassenaar Arrangement) since this clearly states that it will not be used to impede genuine civil transactions. The government hence needs to show that it is serious about making the UK a world leader in electronic commerce by removing all such controls.

Technology Concerns

In many respects, the technology to support digital signatures is immature and not ready to support electronic commerce on a large scale without significant risks. If such risks are carried by consumers, it seems most unlikely that electronic commerce will develop rapidly to accommodate such technologies.

For this reason, it would be preferable at present to limit legislation to technologically neutral measures rather than those intended to provide incentives for the introduction of new technologies. In particular the real value and utility of the technologies for cryptographic digital signatures and certificates is not yet clear and this suggests that legislation that seeks to provide a basis for their introduction is premature.

Acknowledgements

The comments made here are mine alone but they have been much influenced by the many first class contributions made by others. I would like to acknowledge the contributions made by many on the ‘ukcrypto’ mailing list, and especially those of Charles Lindsey and Ian Brown in their recent analysis of the government’s current proposals.

I also wish to acknowledge the direct contribution made by Ross Anderson, Caspar Bowden and (especially) Nicholas Bohm on a number of specific issues covered in this response.

Yours sincerely,

B. R. Gladman

Detailed Comments on “Building Confidence in Electronic Commerce – A Consultation Document” (URN 99/642).

The following comments use the paragraph numbers of the original document.

3. Since the reliability of written signatures is not very high, it is not obvious why signature certification is necessary to achieve such levels of reliability.

11. As the document points out, however, to be compliant with the EU directive such measures have to be “appropriate, effective, necessary and proportionate”.

Footnote 12. It would be better to say ‘can in principle provide’ rather than ‘can provide’ because many of the technical problems of achieving high assurance digital signatures for consumer use have not been solved.

19. In the UK, a person is not bound by a signature that they have not made and it falls to the other party to show that a disputed signature is valid. A rebuttable signature would reverse this situation so that a person whose signature is alleged would have to show that they had not made it. This is a major change and one that the technology currently available cannot effectively support.

20. This paragraph (and the document as a whole) takes an unbalanced view of the technical issues involved in the operation of cryptographic digital signatures.

There are (at least) three issues involved in the management of the keys used to generate such signatures:

Key Generation

Cryptographic digital signatures require the generation of two related keys, one of which (the signature verification key) is published and the other of which (the signing key) has to be under the sole control of the signature owner. If the CA generates these keys, the CA has to be trusted not to keep copies of the secret key since this would allow signature forgery. On the other hand, the average PC cannot provide the safe and secure environment required for user controlled generation of such key pairs.

In consequence, special mechanisms may be needed to generate key pairs and their correct operation will be critical for confidence in cryptographic digital signatures. If such mechanisms are to be trusted there will need to be publicly accountable expert scrutiny of their design, their implementation, their operation and all processes involved in their supply to signature owners. In addition, all aspects of their operation will need to be controlled by such owners. There will also need to be international diversity of supply to guard against subversion of such mechanisms when supplied by some sources.

It is putting the ‘cart before the horse’ to make such mechanisms a subsidiary aspect of signature certification since, unlike certification, they are essential prerequisites for effective cryptographic digital signatures. [the term ‘mechanism’ is used here to describe a hardware, software or hybrid device].

Management of Public Signature Verification Keys

Much of the document is concerned with certification, that is, just one aspect of the management of the public components of signature keys. The document is weak in its coverage of the other issues impacting on the management of signature verification keys..

First, while digital signature certification is a potentially valuable process, it will not always be necessary. In many situations, a signature creation device acceptable to the community as a whole could export signature verification keys directly without the need for a CA. When an owner of such a device is able to pass their verification key directly to those who need it, the avoidance of the certification step will often provide improved signature security.

Secondly, the document places too much emphasis on the application independent aspects of digital signature certification and too little on the importance of application specific knowledge in such processes. For example, the digital signature of a medical doctor (used, for example, to sign prescriptions acceptable to pharmacists) is more likely to be trusted if certified by an appropriate medical authority rather than by an application independent signature certification authority.

Thirdly, the document says little about the need for confidence in signature verification processes. It is not obvious how devices and applications involving signature verification can be built in such a way that the community as a whole can have confidence in them.

Management of Secret Signing Keys

> Signing keys have to be maintained under the sole control of their owners if signature forgery is to be avoided. Moreover, if malicious repudiation by a signature owner is to be prevented, not even the signature owner can have direct access to their signing key.
>
> The secrecy of the signing keys must be maintained during all signing processes and the people making signatures must have confidence that the signing device really does sign what they believe is being signed.
>
> These requirements are considered by many to be beyond the current state of the art. In particular, a typical consumer PC cannot offer the level of security needed for such purposes. Moreover, even if the signing key could be protected, for example, by putting it on a smartcard, this does not overcome the problem of ensuring that documents being signed are those being displayed to signature owners.

Confidence in cryptographic digital signatures hence involves a great deal more than signature certification. The document is currently unbalanced in its coverage of these issues.

21. There is a repeated implication in this paragraph that unlicensed certification authorities will carry extra risk. While this is the objective of the proposed licensing regime, experience with digital signature technologies is currently limited and this could result in the establishment of a licensing regime that fails to tackle the real issues involved. In this situation, unlicensed authorities might tackle some of the critical issues earlier than licensed ones and hence provide more trustworthy services. In particular, if licensed signature certification shifts the burden of proof in signature forgery onto signature owners, then such signatures may actually carry more risk when compared with those provided by other means.

In view of the lack of experience with digital signatures, it is probably premature to consider the introduction of a licensing regime. Moreover, since most early digital signatures are likely to be authorisation certificates designed for use in closed environments, the value of open digital certificates may initially be very limited. It is hence too early to be sure that the benefits of licensing will justify the costs involved. Experience with digital signature certification may eventually show this to be worthwhile but it would be preferable to base legislation on real experience rather than on presumptions about digital signature use that may well prove to be incorrect.

36-38. These paragraphs seem to be based around the concepts of third party confidentiality services whereas the primary confidentiality concerns hindering the development of electronic commerce are ‘second-party’ ones. When companies offer services via electronic networks, consumers want to be sure that these companies will provide adequate protection for any data that consumers supply in the course of transactions. This may involve private personal data but also data, such as credit card numbers, which, if revealed, could create financial risks for consumers. While the confidentiality of communications between traders and consumers is also important, there are already well known ways of achieving such protection without the need for third parties. Hence the government’s aim of promoting electronic commerce would be better served by emphasising ‘second party’ data protection responsibilities rather than the need for third party confidentiality services.

42-45. In practice, it is unlikely to be sensible to tackle liability issues for digital signatures and certificates without considering the domain of application. At one end of the scale, for example, a digital signature may be used to underwrite small value financial transactions where a failure has very limited consequences. Alternatively, however, a medical authority responsible for the certification of signatures for medical doctors could underwrite the signature of a bogus doctor and this could result in the serious injury or even the death of a patient. Liability issues will usually be associated with applications in just this way and this makes it inappropriate to adopt a generic model for handling liability.

It is also worth remembering that many digital certificates will still be useful even if there is no liability when things go wrong. Taking the earlier example of doctor’s signatures, the value of certificates will be more determined by ensuring that the highest possible number are correct rather than by imposing liabilities in the event of error. Although it can be argued that high liabilities will promote certificate quality, this will not necessarily be the case. If, for example, the government were to issue digital certificates to UK citizens for use in obtaining UK government services, the government would be relying on its own certificates and this makes liability meaningless as a quality enhancing measure. Nevertheless, such certificates might be seen by others as trustworthy and might hence be used for identification in just the way that driving licenses are now. Their value here would not necessarily be undermined simply because the government would not compensate third parties if such certificates proved false.

47. As indicated earlier, electronic commerce will depend on applications and products that contain cryptographic sub-elements. Since electronic commerce is international in character, it can only operate effectively if the capabilities needed to sustain it are freely available within an open international market. Export controls on cryptographic products intended for civil use undermine this market and hence seriously hinder the development of electronic commerce. Moreover, such controls contravene the Wassenaar Arrangement (WA) under which they are supposedly justified since this states clearly that it will not be used to 'impede bona fide civil transactions'. Export controls on civil cryptographic products should hence be removed since there can no longer be any serious doubt that they contravene this central provision of the WA. Such a move by the government would provide real evidence of its determination to eliminate constraints on the development of electronic commerce.

48-50. The case that encryption is a threat to Law Enforcement does not appear to be a strong one. In many of the examples quoted the authorities were able to overcome encryption use, albeit with the need for additional resources. Moreover, an international research effort by Professor Dorothy Denning has produced remarkably few examples where law enforcement authorities have been seriously hampered by encryption and even fewer where this barrier has proved insurmountable.

50. The description of 'cryptoviral extortion' in this paragraph is wrong. In this form of attack, a virus encrypts critical information found on a machine using a key available only to the virus writer; it then deletes the original information. The machine owner is then required to meet the virus writer's demands in exchange for the key needed to decrypt and hence recover the critical information deleted by the virus. In this situation prosecution of the virus writer will depend on showing that they alone had access to the decryption key. Far from helping in investigating this crime, police access to this decryption key could undermine successful prosecution by allowing the virus writer's lawyers to claim that this key had been planted by the police in order to frame their client.

53-56. Unless the government intends to prevent the widespread use of encryption, it will not be possible to maintain the effectiveness of intelligence agencies unless their methods of operation change radically. Moreover, since no level of determination on the part of the government will prevent the spread of encryption, law enforcement agencies will also be undermined unless they develop the techniques needed to be effective in an environment where encryption use is ubiquitous.

Law enforcement authorities now need to face the fact that the covert interception of transferred information will progressively become unavailable to them. In consequence, instead of misplaced faith in efforts to prevent or control the use of encryption, they need to invest in the resources and expertise needed to be better than the criminals in this new environment. This is the only solution and to pretend otherwise is a dangerous delusion that will create serious risks for society.

The government should hence consider the creation of a National Computer Forensics Laboratory in order to ensure that law enforcement authorities have access to the expertise they need to operate effectively in cyberspace. With effective accountability provisions, such an organisation would be likely to command the support and respect of both the police and the public.

57-58. The government's announcement that it is reviewing legislation covering interception is welcome. In order not to pre-empt this work it would be sensible to remove these aspects from the current paper (especially so since they have very little to do with electronic commerce).

62. Since the police already have the power to require an individual to produce computerised information in a visible and legible form, it is not obvious why they require additional powers to demand decryption keys.

63. It is reasonable for the police to have the power to require that encrypted information is made available to them in plain text form. They may also need the power to require that the correspondence between encrypted and plain text is demonstrated. They should not have the power to demand decryption keys since this may compromise information to which they have no right of access.

69. The choice of whether to provide plain text or decryption keys should rest with the information owner and not the law enforcement authorities. If necessary, the information owner might be required to show a correspondence between encrypted and plain text.

70. The logic in this paragraph is faulty. It is wrongly suggested that because the government already has encrypted material without a decryption key, the 'evidence' is already in its hands. In other words, all the key does is to make it legible. However, with a perfect encryption system (such as a one-time pad), encrypted data without a decryption key contains absolutely no content information for the corresponding plain text. In such a system, the plain text is dependent on ***both*** the encrypted text and the key and does not exist if only one of these is available. In this situation,

therefore, the government has no information until the key is available to them and this means that supplying the decryption key could certainly be an act of self-incrimination. Although most encryption systems are less than perfect, for practical purposes they behave in exactly this way when a decryption key is necessary for the recovery of plain text (that is, when the plain text cannot be recovered by other means).

72-74. Where information has been obtained under such provisions without the knowledge of the information owner, this owner should be informed about the action taken once enquiries have been completed.

80. As many have pointed out, the notice would need to be served on those who have control of the ***decryption*** process. When a criminal sends a message to an innocent party using public key cryptography, it is the key belonging to the innocent party that is needed for its decryption. It would be grossly unjust to require this key to be divulged. Fortunately, however, most systems employ an intermediate 'session key' of some kind and this can often be made available by the innocent party for specified messages without the need to compromise the secret key of their key pair. This will also allow the correspondence between encrypted and plain texts to be demonstrated. If decryption keys really are needed by law enforcement authorities, such requirements should be limited to short duration session keys.

81-82. There is a widely held view among encryption experts that key-escrow and key recovery techniques carry significant security risks for users, especially so when operated by third parties on a large scale. Given this expert opinion and the lack of evidence to refute it, it is very surprising to see that the UK government is still advocating the use of such approaches. At the very least such advocacy requires that the government provides the evidence that it has used to decide that these expert assessments are mistaken. In the absence of such evidence, UK government advocacy of such techniques amounts to an irresponsible encouragement to users to put their vital information assets at risk.

Annex A (ii). These proposals seem to be built around identity based digital signature certificates. The desire to prevent signature keys validating keys used for confidentiality purposes makes no sense. This is a valid use of signature keys and, if the government now accepts that key escrow is no longer a requirement, there is no reason to continue with this constraint. If this is a licensing condition, it may put licensed authorities at a disadvantage in comparison with their unlicensed equivalents.

Annex A (iii) and (iv). Providers of confidentiality and related services should only be licensed if they can show that the services they provide meet a specified level of security assurance. This does not appear to be a licensing requirement. Moreover, there appears to be no requirements for records or audit trails covering requests for (or the provision of) keys either for owners or for law enforcement authorities. Such records will be essential if such services are ever to be trusted.