What's happened to PETs?

Robin Wilton
Director, Future Identity Ltd

September 27th 2009

Table of Contents

Introduction

Privacy-enhancing Technologies (PETs) have been the subject of research and development activity for some years now, and on the face of it, enough technical elements exist in sufficiently mature forms to allow the deployment of online services which offer substantial advances in privacy protection.

Arguably there has also been a steady rise in public perception of personal data privacy as an issue, fuelled by (among other things):

- growing levels of financial fraud based on various kinds of 'identity theft';
- high-profile instances of mass data compromise;
- work by regulators to publicise privacy issues;
- introduction of breach notification requirements in some jurisdictions.

That said, PETs cannot be said to be mainstream: it is almost impossible to point to any large-scale service deployment which is based on PETs, or differentiated through its use of them. Perhaps more depressing, the R&D activity itself does not yet seem to have resulted in any increased tendency for technology vendors to 'design privacy in' to their products, or for technology adopters to 'design privacy in' to their systems. This paper looks at some of the factors which may help to explain why PETs are still not visible as a mainstream technology, despite their apparently clear ability to help protect the data subjects of personal information.

Research and development activity

It is somewhat invidious to pick out individuals for specific mention, but just as in the field of computer cryptography there are a number of names which figure prominently in any search of the discipline (Turing; Matyas and Meyer; 'Abe' Abraham and Don Coppersmith; Diffie and Hellman; Rivest, Shamir, Adelman and Biham, and so on), so in the privacy field, a number of people show up with greater than average frequency. Here are a few of them -

- The work of David Chaum [1] on privacy, zero-knowledge proofs and 'combining provability with anonymity' stretches back over 25 years (for example, in 1981 he published a paper on "Untraceable email, return addresses and digital pseudonyms", and in 1985 a much-cited paper on "Security without Identification").

- Dr Andreas Pfitzmann [2] of the Technical University of Dresden, similarly, has a 25-year-long research interest in privacy, with a particular focus on linkability as one of the often-undesirable side-effects of systems which are designed without sufficient attention to privacy.

- Dr Jan Camenisch [3] (of the IBM Zurich Research Laboratory) published a paper on privacy-protecting payment systems back in 1994, and more recently is one of the chief architects of IBM's "Idemix" system for anonymous credentials and assertions. In the Spring of 2009 he lectured on "Privacy in the Electronic Society" at the Swiss Federal Technology Institute, looking at technical, legal and regulatory factors in online privacy.

- More recently, the work of Stefan Brands [4] (Microsoft) and Vitaly Shmatikov [5] (University of Texas) has achieved wide visibility in the fields of secure selective attribute disclosures (U-Prove) and the "re-identification" of supposedly anonymised data, respectively.

As well as these individual efforts, there has also been institutional work towards PETs, including projects such as:

- The EU-sponsored PRIME project (2004-2008) and its successor, PrimeLife (current); [6]
- FIDIS – the Future of Identity In Society and its offshoot, the IDIS Journal and workshop series; [7]
- UK Information Commissioner's Office "Privacy By Design" report (Nov 2008) [8]

Technical elements

As well as the theoretical and research work in the field, there also seem to be a number of candidate technologies which are mature enough to contribute to possible PET solutions.

In security/hardware terms, one could point to:

- established encryption techniques
- smart cards and trusted tokens
- increasingly powerful, portable client devices (laptops, palmtops and other mobile devices)

In the software portfolio there are both generic application design elements such as:

- distributed applications
- object-oriented encapsulation (tight coupling of data and methods)
- XML-based messaging (tight coupling of formats and rules)
- web services architectures for distributed data and processing

And more specifically privacy-oriented components such as:

- increasingly granular levels of assertion, isolating identity, entitlements and attributes;
- zero-knowledge proof techniques.

On the face of it, then, the tool-bag is far from empty. However, experience also tells us that complex technology often fails to achieve mass adoption unless it is either completely hidden from the user (e.g. SSL for session-level encryption) or presented to them via easily-understood metaphors which closely match users' mental models of what the technology is intended to do (e.g. Chip and PIN for cryptographically-secured payments).

Arguably, this reliance on metaphors itself introduces further issues relating to user adoption: the more users are hidden from the underlying complexity of the technology they use, the more they have to rely on the accuracy of those metaphors and the consistency with which they represent what is actually happening – as opposed to being able to rely on a direct and accurate understanding. Cryptographic technologies seem to have become steadily less and less intuitive as they have progressed from symmetric to public key to elliptic curve algorithms, and there is really no prospect of the average user being able to determine whether the technology is actually doing what it claims, and with the expected level of security.

Even the technically competent can get caught out – as was illustrated last year, when a combination of DNS cache poisoning and a flawed cryptographic key generation module resulted in potentially devastating failures in implementations of the OpenID authentication protocol. There is a brief description of that incident on my former blog, here [9].

End-user motivation as a factor

Back in early 2006 I expressed the view that "Privacy is the new Green" - by which I meant that I thought advocates of a stronger approach to personal data privacy were regarded with the same bemused semi-tolerance as the 'tree-hugging eco-warriors' of preceding years. Ecological and sustainability issues are now, of course, squarely in the mainstream of both policy and public awareness, but privacy concerns are often still treated as if they were a niche lobbying interest, compelling only to conspiracy theorists, and those who, it is implied, must have "something to hide".

All else being equal (as if that were ever a realistic precondition), the change from mainstream to egregious behaviour in any area of life usually involves a tipping of some balance between preference for one option and preference for another. By analogy, what does this suggest for the adoption of PETs?

1 - LPG analogy:

Someone dependent on LPG outlets cannot drive around with the same abandon and spontaneity as someone who can rely on the petrol distribution infrastructure. Indeed, there may even be places it is unrealistic for them to expect to visit.

If few or no services are available in PE forms, it may simply not be viable to expect to 'get by' using only privacy-respecting service providers. That may only be an inconvenience if it's simply a matter of buying books from one or other online retailer – or getting them the old-fashioned way, from a book-shop. However, for non-discretionary services such as benefit claims or healthcare, the implications may be far less trivial.

Perhaps more significant: there is a comparison to be drawn between the characteristics of the LPG infrastructure (comparatively sparse as a proportion of the fuel network as a whole, and offering a reduced choice of outlets and suppliers), and some nascent candidate PETs such as the TOR network [10]. The work of Steven Murdoch and Piotr Zielinski [11] appears to show that despite the lengths TOR goes to to protect users from traffic analysis, traffic data can be de-anonymised to a surprising degree even if an attacker has access to only a small proportion of it. In other words, some forms of PET require a certain critical mass if they are to be effective: and below that critical mass, may in practice offer a false impression of privacy protection.

2 - Organic food analogy:

Here, the user may be presented with a small, but nonetheless viable number of options – but may be obliged to compromise in other respects:

1. a limited choice of suppliers;
2. a restricted choice of products (one kind of apple rather than several; or bananas but no tangerines);
3. increased cost.

In privacy terms, these may translate to

1. increased linkability of transactions and therefore behaviour;
2. only partial mitigation of privacy risk (i.e. If some service providers are privacy-friendly and others are not). In some instances that may be acceptable: for example, even if only some websites can process assertions of the form "Is over 18", that may still be an improvement on having to disclose my date of birth everywhere. On the other hand, in use-cases such as the protection of victims of abuse, vital witnesses, under-cover police officers and so on, "80% pseudonymity is not really good enough".

What should PETs be helping us to do?

It is interesting to note the way in which some existing technical building-blocks have been 're-branded' as privacy-enhancing – for instance, encryption of 'data at rest' is described as privacy-enhancing because it helps mitigate the potential impact of a data breach. Similarly, I have heard the same view expressed for 'data in motion': specifically, that 'if the HMRC disks had been encrypted before transmission, the privacy of those concerned would have been protected'. Up to a point, I agree: clearly, one threat to privacy is that of personal data being inappropriately accessed, either while it is being stored by a legitimate data custodian, or while it is being shared with one.

However, I would argue that the portrait this paints of privacy risk is neither complete nor, consequently, useful. For instance, if I keep all my personal data to myself (storing it on disk and encrypting it), that may satisfy some definitions of privacy, but it does not help us understand how I can preserve my privacy while at the same time leading a "normal" online existence – browsing the Internet, conducting e-commerce transactions, sending and receiving emails, participating in social networks, and making use of online public sector services.

To be useful, a definition would need to explain how I can *disclose* my data to those I interact with online, while still preserving my personal privacy. In this sense, three criteria seem useful in assessing the privacy-enhancing qualities of a given implementation:

1. preservation of contextual integrity,
2. protection of privacy beyond first disclosure,
3. meaningful consent and control.

Preservation of contextual integrity

An inspection of personal data suggests that one way to analyse it is in terms of the context in which it is used and disclosed. For example, there is a subset of my personal data which is fairly clearly related to my activities as a driver: I have specific credentials (a driving licence), which record both generic data about my identity, and sector-specific data about which vehicle types I am entitled to drive, whether or not I have any endorsements, and so on. Elsewhere, there is a subset of my personal data which is fairly clearly related to my healthcare; I have a health-care specific credential (an NHS patient number), and a set of patient records about conditions I have, treatment I have had, and so on.

To be sure, there are cases in which data about my health can correctly appear in the records about my activities as a driver: if I am registered blind, prone to epileptic seizures, or subject to some other condition which materially affects my ability to drive, I might expect that to be reflected not just in my health records, but also in those relating to my driving licence.

However, if I were to find that my driver details also contain a record that I have had my appendix out,

or that I have seen my GP about hay fever, I might feel that my privacy has been violated – because those disclosures would indicate that the contextual integrity of my personal data has not been respected. In other words, I tell my doctor some things because of my relationship with her as my doctor, and I tell the DVLA some things because of my relationship with it as a driver. Those contexts make it appropriate for each of those parties to know some things about me but not to know others.

Thus, it seems to me that a large part of preserving my personal data privacy comes down to ensuring that the contextual integrity of may disclosures is respected. This gives rise to a second set of requirements which privacy-enhancing technologies might be expected to help satisfy.

Privacy protection "beyond first disclosure"

As suggested above, the "paradox of privacy" is that it consists not in keeping all my data to myself, but in sharing it – but doing so with knowledge of, and control over, the contextual factors which apply (including appropriate and informed consent). Let's return to the example of encrypted 'data in motion', substituting the following actors for the more canonical Bob and Alice:

- Primrose wishes to be able to disclose data while preserving her privacy;
- Reece is the intended recipient of Primrose's disclosures;
- Malcolm is a malicious party who wants to access Primrose's personal data.

If Primrose wants to share data with Reece, without Malcolm being able to intercept it en route, Primrose might choose to encrypt the data. So far so good – if Malcolm intercepts the message, it may be impractical for him to try and decrypt it. However, Primrose does want Reece to be able to read the data, so she has to give him the means to decrypt it; if Malcolm can get to the data now, it will be in clear... and this might happen because Malcolm can get at Reece's system, or because Reece (or someone else with access) writes the data down and discloses it to Malcolm. As soon as Reece has decrypted the message, the *technical* protection which Primrose applied has been nullified. If Primrose's privacy is to be preserved beyond first disclosure, that protection must be based on other things (such as a confidentiality agreement between Primrose and Reece, effective access control to the data in clear on Reece's system, and so on).

Note also that encryption of the 'data at rest' on Reece's system after receipt provides only partial privacy protection for Primrose: to be useful, the data has to have been in clear for Reece to read, and if at that point Reece simply writes it down, encryption of the copy on disk does not prevent Reece from disclosing the information.

Put another way, data encryption does not provide effective privacy protection 'beyond first disclosure'... and as I argue that the idea of privacy without disclosure is not a useful one, I have to conclude that on its own, data encryption is not a very effective privacy-enhancing measure.

Meaningful consent and control

If "contextual integrity" is a desired end and "protection beyond first disclosure" suggests a set of requirements which go towards satisfying it, we should also consider how PETs might make it possible for disclosures to be subject to appropriate levels of informed, rational consent and auditable control.

A UK Technology Strategy Board-sponsored project, VOME (Visualisation and Other Methods of Expression) [12] has the goal of looking at how users conceptualise their own privacy – and thus of building a better picture of how users' privacy-related decision-making can be improved to lead to

better privacy outcomes.

For instance, a data subject's consent (unless it is simply considered to be implied) is of little practical use if its giving is not recorded, and if its revocation is not possible... and yet serious research into how this can be achieved is only recently getting under way – in the UK at least – through projects such as EnCoRe (Ensuring Consent and Revocation) [13].

What's more, in the absence of workable mechanisms for protection beyond first disclosure, it's debatable whether users can be guaranteed a meaningful level of technical control over the subsequent use of data they disclose. Two areas of work here are

a) privacy preference expression/enforcement and
b) "sticky policy"

Privacy Preference Expression/Enforcement

Very briefly, privacy preference expression languages (PPELs) are an attempt to formalise the expression and subsequent enforcement of a user's privacy preferences. One problem here is that matching what a data subject wants to allow and what a data consumer wants to do is usually more of a semantic exercise than a syntactic one. To give an example: data protection law in most countries establishes the idea that if data is collected for a purpose, it should not subsequently be used for something else. On the face of it, then, all a PPEL ought to have to do is compare the stated purpose of collection with the stated purpose of use and make a decision as to whether they are compatible.

This raises one semantic problem, one pragmatic problem and one cultural one.

- The semantic problem is that statements of 'purpose of collection' are often extremely broadly worded (think of the telephone help-line disclaimer "your call may be recorded for quality control purposes"...), while statements of 'purpose of use' are usually much more specific ("correct addressing of a marketing mail-shot"). It can be very difficult to automate the matching of the two in such a way as to get robust and reliable results.

- The pragmatic problem is this: if an organisation perceives that it is being inhibited from fulfilling its primary purpose (whether that is collecting taxes, treating patients or selling DVDs) because an automated privacy preference matching engine is saying "no" too often, the pressure to relax the matching engine's rules – if not turn it off altogether – will become irresistible.

- The cultural problem is closely related to the pragmatic one, and was illustrated by the HRMC data breach case. In that instance, the risk of disclosing sensitive personal data inappropriately and on a mass scale was assessed against the cost of spending £5,000 to extract the few hundred records which were actually requested... and the equation came out in favour of massive disclosure. The only answer to the pragmatic problem is for the value of personal data to be differently quantified, and the risk arising from it to be differently assessed. As long as personal data privacy is given the same importance as in the HRMC instance, organisations will always opt to ignore or disable privacy protection rather than incur additional cost or impair what they see as their over-riding goal.

A white paper, published by the Liberty Alliance in 2003, sets out the principles of Privacy Preference Expression Languages in a technology-neutral way, but illustrated with reference to existing technical

specifications such as ID-FF, ID-WSF, SOAP and P3P. [14]

"Sticky Policy"

Sticky policy is related to the ideas of privacy preference expression and privacy beyond first disclosure. It consists in finding ways to bind a user's privacy preference statements effectively to the data in question, in such a way that any recipient of the data cannot either claim to have been unaware of the preferences, or (ideally) disregard them when processing the data disclosed. Achieving 'sticky policy' requires a combination of technical mechanisms and other disciplines such as reliable audit and logging.

One approach is suggested by the research work of Radia Perlman at Sun Microsystems. Her "Ephemerizer" design [15] is based on the idea that Primrose, when disclosing her data, encrypts it, but instead of distributing the corresponding key to Reece, Primrose instead lodges it with Polly – who runs a policy-enforcing "Ephemerizer" service. Reece must apply to Polly for the key in order to open the disclosure, and this provides Polly with an opportunity to ensure that Reece signs up to Primrose's privacy preferences, log Reeces' acceptance, and enforce time or frequency limits on Reece's ability to access the key in question, and hence also the data it protects.

There are still some issues with this approach. One is that it is, to some extent, vulnerable to the same criticism as encryption of data in motion: once Reece has legitimately opened the disclosure for the first time, there may be nothing to prevent him from simply writing the information down – thus bypassing all the technical protections. Primrose may at least be able to claim that if her data is subsequently found 'in the wild' and stripped of its technical protections, Reece may have violated the privacy terms to which he (auditably) signed up... but that assumes, of course, that Reece is provably the only person to which that disclosure has ever been made, and for many disclosures that is not a trivial exercise. For example, it may not offer any practical protection for common items of personal information such as Primrose's address.

Another issue with this approach is that Primrose has to have a very trusting relationship with Polly. Polly's ability to build up a very detailed picture of Primrose's online behaviour might, under the wrong circumstances, be just as damaging to Primrose's privacy as unmanaged disclosures.

One factor in favour of the "Ephemerizer" approach is that it can be constructed from existing, mature and well-understood technical components and management techniques. Encryption, digital signing, and audit/logging are all well-established disciplines, though the key management disciplines required qualify, unfortunately, for that description of PKI as "a technology with the longest take-off run ever".

There are two alternatives which promise to overcome some of the linkability and aggregation issues described here, namely the Idemix and U-Prove designs – but these are both early enough in their lifecycle that they are yet to be visible in live implementations as opposed to presentation or demo form.

Frameworks for assessing adoption factors

It may indeed be that the candidate technologies mentioned so far are both viable and effective; however, experience clearly indicates that good technology is not enough. The ICO's Privacy By Design report [8] sets out a number of other factors which, if missing, inhibit an efficient and workable market in any new technology or discipline. Among these are:

- shared standards;
- experience of deployment and management;
- common processes;
- like-for-like comparability;
- accreditation standards;

These fit into a broader contextual picture including:

- conceptual framework
- legal context
- implementation readiness
- process readiness
- regulatory and compliance criteria
- adoption culture/behaviour

Failure in some of these areas is merely inhibitory; in others it is fatal. And, of course, there are multiple stakeholders who must all feel that their interests are being met if adoption is to succeed. Not surprisingly, given the nuanced and variable nature of privacy as a concept, the picture we get is of a complex ecosystem of inter-related factors and influences, in which the failure of one or more components may render the whole unappetising or unworkable.

Up to now in this paper, I have focussed primarily on the technical factors. However, even the best technology is unlikely to be able to survive the failure of some of the other elements. In fact, the history of computing is littered with dead technologies, killed off by competitors which were technically inferior but did a better job of satisfying other adoption criteria... at the right moment. At the risk of sparking a techno-jihad, I might cite as examples:

- IBM token ring LAN vs Ethernet
- Sony Elcaset vs standard audio cassette
- CISC vs RISC computing
- Bipolar vs CMOS chip technology...

Even those who profoundly disagree with my choices probably have pet (no pun intended) alternatives of their own.

I therefore wanted to look at whether there are any models which might make it simpler to analyse and perhaps even influence this complex ecosystem.

The Question of Timing

It's a truism that technical innovations often happen "before their time". Sir Clive Sinclair's electric-engined C5 (launched in January 1985) was a commercial flop at the time, but technically it is a genetic ancestor of the petrol-electric hybrids which may today offer a more sustainable future in personal transport. The following diagram illustrates the principle of 'innovation peaking earlier than adoption readiness'.

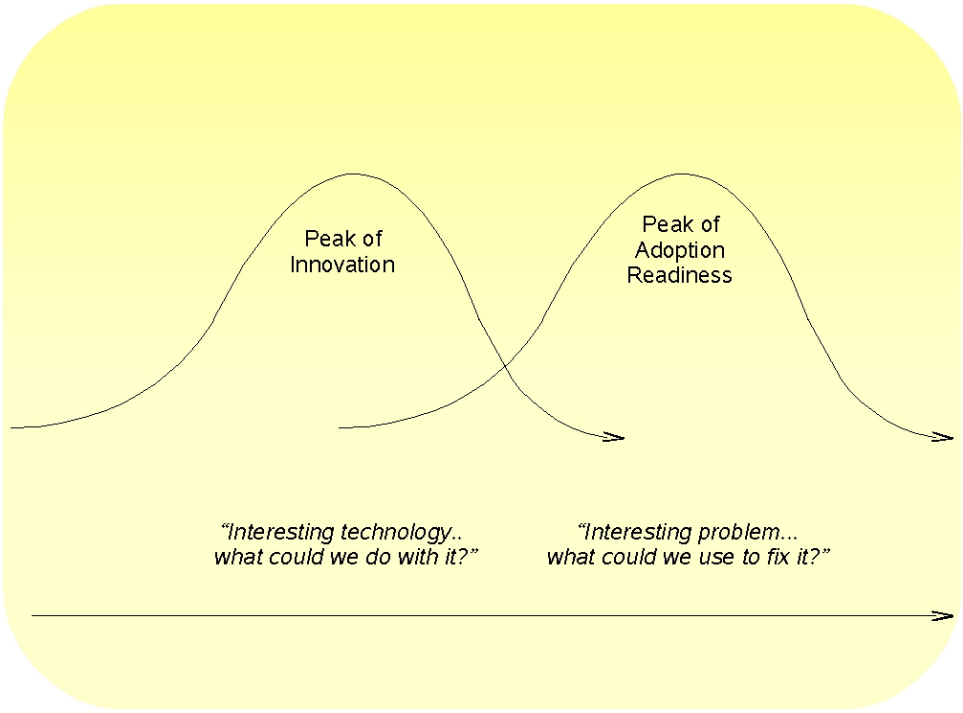

Illustration 1: Timing

The snag, of course, is that conceptual innovation has a shelf life – and there are probably many (by definition unrecorded) instances where the 'interesting problem' has been perceived by someone who does not or cannot identify the right candidate conceptual innovation at the right time.

Web-based "Social Networking" as a counter-example

In this respect, current patterns of online service deployment may offer a counter-example. While I instinctively dislike the term Social Networking (and if you're interested, there's a short blog post here where I say why [16]), the evolution of services such as Twitter suggests that sometimes, the innovation and adoption curves can sometimes not only follow smoothly, but can then give rise to further waves of innovation. The developers of Twitter probably had at least some idea of what if might be 'for' – but I'm sure they did not (and could not have) forecast all the various purposes to which users now put it, including:

- gossip (OK, they probably did predict that one)
- sharing web content
- keeping up with the news media
- political protest
- organising spontaneous gatherings, and even

- checking when the International Space Station will next pass overhead... honest. [17]

The point is that it's increasingly viable to put online services into the public domain without knowing what they will end up being used for. So, in this instance, the initial innovation was followed by an adoption curve, which in turn gave rise to waves of further innovation as a general-purpose service platform is put to new and unforeseen uses.

The "S-curve" model

As one of the outputs from the PRIME project [6], John Borking and others cited the previous work of Richard Nolan [18], Watts Humphreys [19] and Everett Rogers [20] in describing the application of an "S-curve" model to the analysis of innovation adoption. Apparently Rogers' work was foreshadowed as early as 1903 by the French sociologist Gabriel Tarde's observation that cultural diffusion often follows an s-shaped curve [21]. Borking recaps the PRIME work's findings in this article [22].

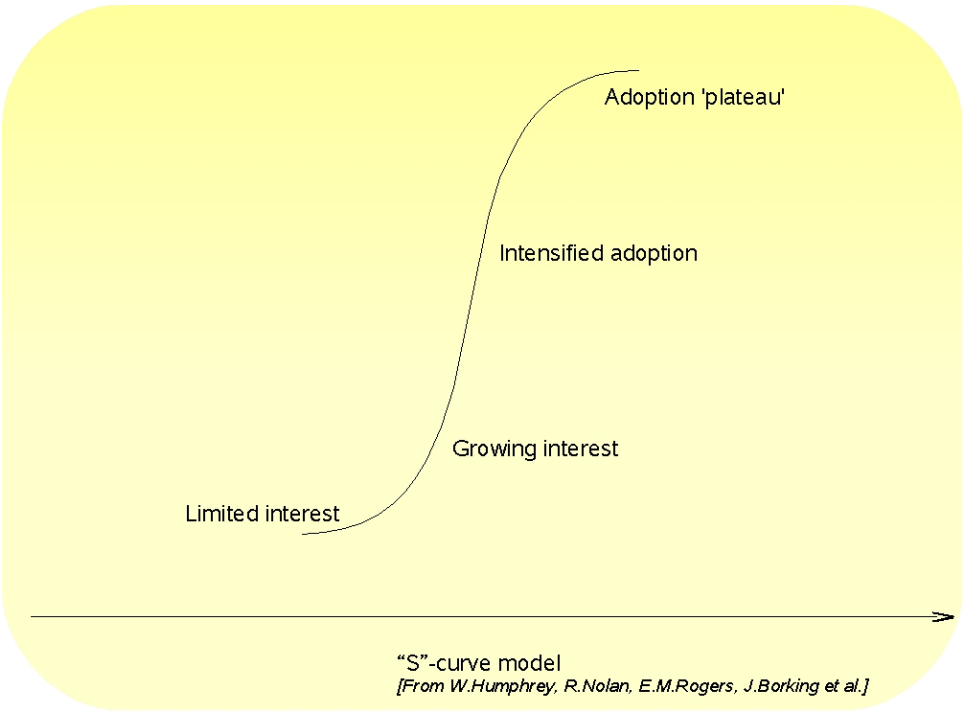

Illustration 2: S-curve model

In summary, one way to analyse cultural diffusion (and therefore, by implication, awareness, understanding and adoption of new technology) is to track its progress through the phases of this S-curve.

However, presented in this simplified form, the S-curve is probably most useful as a retrospective tool for charting how far along the curve any given innovation reached. It is less useful as a means of predicting likely adoption factors and thus being in a position to influence them positively.

What I suggest, therefore, based in part on Borking's conclusions, is that the S-curve is in fact a composite, and that each phase actually consists of an equivalent 'sub-curve' which is subject to its own set of adoption factors. This is illustrated in the diagram below:

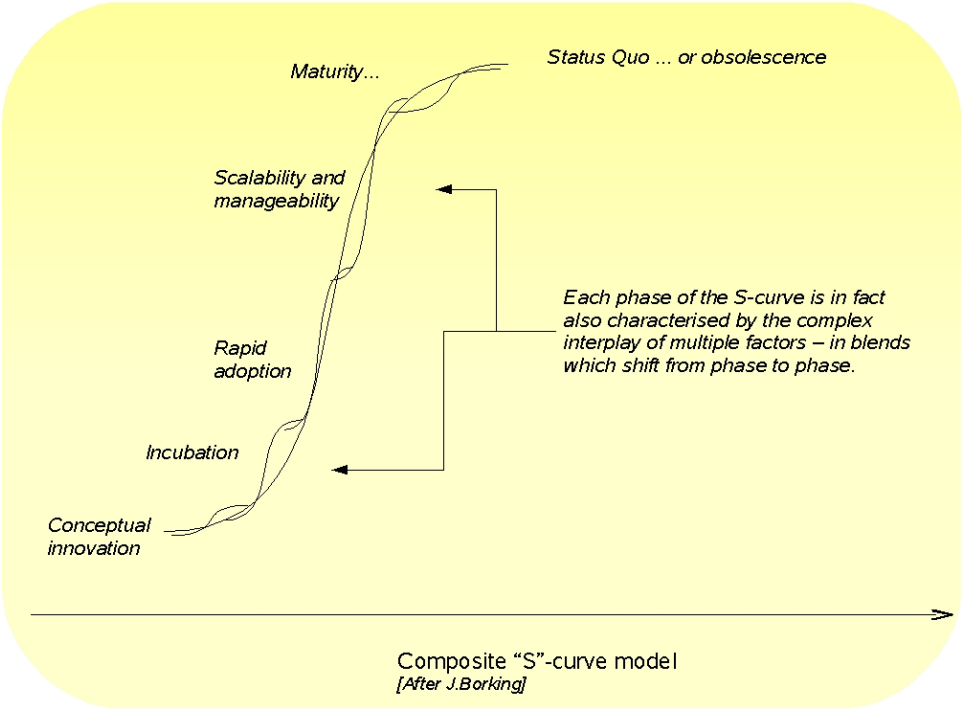

Illustration 3: Composite S-curve

Here, I have given some rough examples (to the left of the curve) of the kinds of phase through which technical innovations pass. At each stage there will be not just one, but a set of critical success factors to satisfy. To be sure, some of them (such as legal frameworks) are persistent – but what this diagram illustrates is that the blend of success factors will change from phase to phase: it's not just a matter of overcoming a given set of adoption obstacles – you have to go on adapting, and solving successive sets of adoption problems as the S-curve gives rise to them.

In other words, it's no good jumping your hurdles before you come to them; correct timing is every bit as important as flawless technique.

Preliminary conclusions

The first test of this 'composite S-curve' model must be – does it offer useful guidance when applied to a set of the adoption/readiness factors identified earlier. With that in mind, let's revisit the list of contextual factors I cited earlier, and 'plot' them against the phases of the composite S-curve. Clearly, one could do this in the form of a matrix for completeness' sake, charting the relevance of each adoption factor at each stage of the S-curve – but for the purposes of this paper, I will restrict myself to discussing each adoption factor in the context of the S-curve phase to which it seems (subjectively) most relevant. Other analyses may apply a different priority.

1. conceptual framework
2. legal context
3. implementation readiness
4. process readiness
5. regulatory and compliance criteria
6. adoption culture/behaviour

1. Conceptual Framework (Conceptual Innovation phase)

At this stage, a conceptual framework usually needs to exist; the "giant leap forward" kind of innovative discontinuity is rare, compared to incremental inventions based on the fruits of recent previous progress. And this is visible in the identity/privacy world: the solutions to problems of distributed online identity will come from progress in federated identity management, which in turn arose from innovations in enterprise-centric identity.

For the innovation to take root, there must also be a conceptual framework which allows different stakeholders to share their perspectives on the innovation in question. In the course of a two-year series of privacy round-tables (2007-2008) I found that one of the greatest obstacles was the lack of a shared conceptual framework and vocabulary which would allow participants from different backgrounds to contribute constructively to the debate [23].

Developing a set of simple models to overcome this obstacle, we found that subsequent multi-party debates made more progress, more rapidly.

2. Legal Context (Incubation phase)

Except in the realm of criminal innovation (and we shouldn't write that off... criminal problem-solving is every bit as creative and effective as the law-abiding kind), innovation needs to take account of whether or not there is applicable law. If there isn't, that doesn't necessarily mean that the innovation cannot or should not happen. In fact, there are plenty of examples of technological progress outstripping the capacity of the law to sustain suitable governance: software patenting and the copyright protection of digital media are two illustrative examples.

Nevertheless, if an example of innovation either runs contrary to existing law or promises to take adopters into uncharted legal territory, it is unlikely to benefit from conventional support in the early, low- to no-adoption phase.

What is evident currently in the privacy world is that the applicable law is confused. Definitions of key terms such as 'privacy', 'personal data' and 'personally identifiable information' differ from one

regulatory environment to another, and in some cases are inconsistent. The relationship between personally identifiable information, identity, privacy and data sharing is not codified, and existing legal definitions sometimes fit badly (or not at all) with the technical paradigms of Web 2.0, 'cloud computing' and 'Vendor Relationship Management'.

(As one brief example: when a user voluntarily consents to have their data distributed under the autonomous control of a cloud-based storage algorithm, and therefore has no idea in which national jurisdiction some of the data will end up, how can existing (nation-centric) notions of 'data controller' be applied?).

3. Implementation readiness (Incubation phase)

Technically, some elements of effective privacy enhancement are likely to depend on options which are currently immature at best. For example, the appropriateness of a disclosure of personal data is highly contextual; context management, in this sense, will depend not just on the data itself, but on a range of contextual factors which would have to be represented as metadata, if at all. There are currently few standards for such metadata and even fewer examples of technology to process and act on it.

4. Process readiness (Rapid Adoption phase)

As mentioned above, effective privacy management implies an advanced ability to manage contextual metadata. Organisations are still generally poorly prepared to do this; for instance, even among that subset of organisations which effectively maintain an inventory of the personal information they collect and process, a much smaller subset is equally capable of managing the associated meta-data. Similarly, few organisations have effective means in place to allow end users to express meaningful privacy preferences, or to enforce such preferences if they can be expressed.

5. Regulatory and compliance criteria (Scalability and Manageability phase)

As the ICO "Privacy By Design" paper notes, there can sometimes be little or no connection between compliance goals and privacy goals. In other words, a regulated organisation may meet current compliance criteria while still doing business in a privacy-hostile way. There is also a lack of common accreditation standards for compliance assessors in the privacy domain, even in individual regulatory environments – and for every national regulatory question, the global internet poses many more cross-border ones. As mentioned above, the ability of technology to stretch current legislation (as in the examples of software patenting and digital copyright) means that mass adoption can rapidly create legal 'blind spots' on a massive scale.

6. Conceptual framework (Maturity phase)

In John Borking's paper, he notes that the end of each S-curve often represents just a transition/decision point before the start of a new curve. And so it is in this case. If we assume that PETs do ultimately achieve mass adoption, we must also assume that by that stage, there will have been a qualitative change in the way end users tend to view their online privacy. That, in turn, suggests that they will have developed (consciously or not) a new conceptual framework for doing so. And what next? That point almost certainly signals the start of the evolution of a (new) new conceptual framework, and the beginning of another S-curve.

References

[1] Dr. David Chaum; list of publications. http://www.chaum.com/articles/list_of_articles.htm
[2] Dr. Andreas Pfitzmann; list of publications since 2002. http://www.inf.tu-dresden.de/index.php?node_id=703&ln=en
[3] Dr. Jan Camenisch; selected publications. http://www.zurich.ibm.com/~jca/publications.html
[4] Dr. Stefan Brands; publications from 1993-2007. http://www.informatik.uni-trier.de/~ley/db/indices/a-tree/b/Brands:Stefan.html
[5] Dr. Vitaly Shmatikov; list of publications since 1998. http://www.cs.utexas.edu/~shmat/index.html#pub

[6] EU PRIME and Primelife projects https://www.prime-project.eu/ , http://www.primelife.eu/
[7] EU FIDIS and IDIS programmes http://www.fidis.net/resources/deliverables/other/
[8] UK Information Commissioner's Office report; "Privacy By Design" (Nov 2008) http://www.ico.gov.uk/upload/documents/pdb_report_html/index.html

[9] OpenID, DNS and Debian – the perfect storm http://blogs.sun.com/racingsnake/entry/one_factor_trust_multi_factor
[10] TOR Project website http://www.torproject.org/
[11] Steven Murdoch and Piotr Zieliński; Sampled Traffic Analysis by Internet-Exchange-Level Adversaries: http://petsymposium.org/2007/papers/PET2007_preproc_Sampled_traffic.pdf

[12] VOME project (Visualisation and Other Methods of Expression) http://www.vome.org.uk/
[13] EnCoRe project (Ensuring Consent and Revocation) http://www.encore-project.info/

[14] Liberty Alliance; Privacy Preference Expression Languages (2003) http://www.projectliberty.org/liberty/resource_center/papers/privacy_preference_expression_languages_whitepaper_pdf
[15]Radia Perlman: "The Ephemerizer: Making Data Disappear" (2005) http://research.sun.com/techrep/2005/smli_tr-2005-140.pdf
[16] Why we should distrust the phrase "Social Networking" http://futureidentity.blogspot.com/2009/07/pointer-to-tech-and-law-blog.html

[17] Twitter and the International Space Station http://www.twisst.nl/

[18] Prof. Richard Nolan; "Stages of Growth" model (1973-1979) http://en.wikipedia.org/wiki/Stages-of-growth_model
[19] Watts Humphreys; "Capability Maturity Model" (1989) http://en.wikipedia.org/wiki/Capability_Maturity_Model
[20] Everett Rogers; "Diffusion of Innovation" (1995-2003) http://www.cw.utwente.nl/theorieenoverzicht/Levels%20of%20theories/macro/Diffusion%20of%20Innovation%20Theory.doc/
[21] Gabriel Tarde; "Diffusion of Innovation" (1903) http://en.wikipedia.org/wiki/Diffusion_of_innovations
[22] John Borking; "Organizational Motives for Adopting PETs" (2009) http://tinyarro.ws/Borking-PETs

[23] Liberty Alliance Privacy Summit Programme – London-Basel Summit Report http://www.projectliberty.org/liberty/public_community/privacy_summits