On technology neutral policies for e-identity: A critical reflection based on UK identity policy

Edgar A. Whitley *

Abstract: This paper reviews the arguments for technology neutral e-identity policies. It uses the recent experience of identity policy in the UK, as well as a consideration of technological developments, to distinguish between two perspectives on technology neutral policies: legal and technological. Whilst the legal perspective on technology neutrality is intended to provide legal certainty, it fails to address discontinuous technological developments such as zero-knowledge systems and risk based assessments of identity and attribute claims. These are transforming the basis of identity policies and highlight the challenges of proposing technology neutral identity policies in law. The paper then applies the technological critique of technology neutrality to review a recent study on identity, authentication and signature policy in the EU.

1. Introduction

For apparently intuitive reasons, many EU regulations and policies in the area of technology and communications are intended to be “technology neutral”. They typically require that “national regulatory authorities take the utmost account of the desirability of making regulation technology neutral, that is to say that it neither imposes nor discriminates in favour of the use of a particular type of technology” (EU, 2002 para 18). It is often suggested that this is because policies should be based on general, context-free principles rather than instantiations of technology.

Focussing on specific technologies carries the risk that the technology might change rapidly or become obsolete, rendering the associated regulations ineffective and requiring special parliamentary time to refresh them. For example, just as wiretap regulations have had to develop and be redefined as interception capabilities moved from alligator clips on copper wires to IP packet sniffing (Diffie & Landau, 2009) so regulations on e-mail interception which are based around inspecting SMTP packets are problematic in situations where browser-based web-mail predominates (Whitley & Hosein, 2005) or lawful access to communications context data is highly dependent on what counts as context and what is effectively content (Escudero-Pascual & Hosein, 2004).

These experiences from the regulation of technology resonate with recent social science thinking on science and technology that questions the extent to which technology can meaningfully be seen as a distinct area of activity that is separate from social activity (Latour, 1993; Orlikowski, 2010; Suchman, 2007). Instead, these scholars seek to understand “how matter comes to matter” (Barad, 2003; Orlikowski & Scott, 2008), returning the “missing masses” of technology to the policy debate (Latour, 1992) and taking the specifics of technology seriously (Martin & Whitley, 2012).

These concerns are particularly important because it is increasingly recognised that the technological itself embodies political choices. Langdon Winner’s classic paper “Do artifacts have politics?” (1980) illustrates this with perhaps the most famous example he presents being the “extraordinarily low” overpasses on the parkways on Long Island, New York. Winner suggests these were “deliberately designed” by Robert Moses, “to specifications that would discourage the presence of buses on his parkways” leaving the parkways (and beaches) free for the “automobile owning whites of ‘upper’ and ‘comfortable middle’ classes” (pp. 123–124). (Although see, Woolgar and Cooper (1999).)

In some cases, political and social choices may be explicitly designed into regulations about technology to encourage or discourage their uptake. As an example, the UK has recently introduced a “digital by default” policy, the intention of which is both to provide better services for citizens and to deliver gross annual savings of more than £1.3 billion (Cabinet Office, 2010). Winner’s analysis, however, warns us that some political decisions may be hidden within apparently neutral propositions, such as the design of parkways. In other cases, the political consequences may be entirely autonomous of individual decision-makers (Winner, 1977) but still need appropriate oversight. In this context, claims that a policy is technology neutral and based around general principles need to be treated with suspicion.
On technology neutral policies for e–identity: A critical reflection based on UK identity policy

as it might, in fact, reveal precisely the kinds of politics that Winner identified. e–identity policy is one such area where claims for technology neutrality are frequently made.

Recent UK experiences around identity policies will be used to explore potential limitations of technology neutral policies. The unique features of the UK case make explicit a range of technological assumptions that are hidden in more conventional or less problematic national identity policies (e.g. see discussion of implicit assumptions in Cormack, 2012) and provide a useful basis for evaluating the scope and applicability of technology neutral policies and highlight a distinctive technological viewpoint on this issue, a viewpoint that extends the existing legal discourse surrounding technology neutral policies. The paper then analyses a recent study on identity and authentication services (IAS) (IAS Project, 2011b) from this perspective.

The paper begins by reviewing the recent UK identity policy, starting with the recently abandoned National Identity Scheme, as legislated for in the Identity Cards Act 2006. It was claimed that the Identity Cards Act was technology neutral, ‘enabling’ legislation that would not restrict the development of identity policy. However, it will be argued that key design choices were hard–coded into the legislation. The UK Coalition Government’s alternative Identity Assurance programme will then be briefly presented to illustrate how many similar goals can be achieved in very different ways. The paper next reviews the legal arguments used to support technology neutral policy based around general, context–free principles. This is followed by technological consideration of the same issue including some recent examples of technology–based innovations that offer privacy–friendly identity solutions that do not fit within the existing general principles of e–identity policy. The paper then uses this perspective to assess critically the IAS proposals before ending with a discussion of the implications of this perspective on e–identity policy more generally.

2. The UK Identity Cards Act: Enabling legislation or wiring in design choices?

In 2005, the Labour Government presented plans to introduce identity cards to the UK for the first time since the Second World War (Agar, 2005). The National Identity Scheme (“the Scheme”) would include the use biometrics and would be based on a centralised National Identity Register (“the Register”) containing the identity details of all UK citizens and residents as well as an audit trail detailing whenever identity claims were verified against the Register. Whilst this audit trail would be useful to citizens seeking to query particular transactions, it also opened up the possibility of surreptitious surveillance of citizens by state agencies (e.g. Lyon, 2009).

To support the Scheme, the government introduced primary legislation in the form of the Identity Cards Bill, which, after a controversial passage through Parliament (see Whitley & Hosein, 2010a) became law in 2006 (Wadham et al., 2006).

Throughout the Parliamentary debate about the legislation Home Office Ministers repeatedly emphasized the fact that the Bill was ‘enabling legislation’ that would “allow” a National Identity System based on identity cards to be introduced. As a result, they stated that there was “much still to be done in terms of detail, regulations and all the other elements” [Tony McNulty, 28 June 2005 : Column 1253].

Many of the details of the Scheme were not included in the Act, with these details being left to secondary legislation and statutory instruments. The use of secondary legislation is not without its critics as, in practice, the debates about them are often poorly attended and so effective scrutiny of the details of the Scheme could be limited, raising the prospect of what Conservative MP Edward Garnier described as “legislation by statutory instrument” [18 October 2005 : Column 804].

The absence of such important technological details in “enabling legislation” makes it particularly difficult for Parliament “to scrutinise the proposed measures effectively” (Constitution Committee, 2009 Recommendation Paragraph 474).

Regardless of these practical considerations, one of the main arguments in favour of “enabling legislation” is that it allows for a “technology neutral” policy as “it may be counter–productive to adjust current regulation incrementally” (Lusoli & Compañó, 2010). Rather than specifying in legislation what technological measures might need to be put in place, this form of legislation allows for these details to be added at a later stage, such as during the procurement process.

Nevertheless, achieving the appropriate level of non–specific detail can be problematic. For example, the final version of the Act states that an individual may be required to allow “his fingerprints, and other biometric information about himself, to be taken and recorded” thus both leaving some details to be added
at the procurement stage (will face and iris biometrics be used?) whilst also specifying that fingerprint biometrics will be used by the Scheme.

The Act further confuses the distinction between technology neutral legislation and legislation with specific design implications in the role of the Register. Thus, whilst the Act does not completely specify all of the biometrics to be stored by Government, it does specify that the Secretary of State “establish and maintain a register of individuals” (s.1(2)) that includes “information about occasions on which information recorded about him in the Register has been provided to any person” (s.1(5)(i)) (i.e. the audit trail mentioned above). Schedule 1 (6) of the Act also specifies other audit details that are recorded on the Register.

This is a very detailed (political) design specification for the Scheme and its operation. Whilst nominally neutral about the technology it actually implies a very particular way in which the Scheme would be used in practice—one that enhances the powers of the state and downplays the rights of privacy for citizens. For example, it strongly suggests identity verification based on a centralising approach to identity management that involves verifying details such as biometrics against those stored on the Register when confirming someone’s identity. This process would also involve creating an associated audit trail record.

In contrast, the legislation appears to rule out formal use of identity verification simply against details held on the card, with no audit trail record created. The reasons for this design decision have never been disclosed however an internal “benefits overview” document (Home Office, 2005) noted that the Scheme (and particularly, the centralised database of fingerprint biometrics stored on the Register) would improve the ability of police to detect crime by “increasing the likelihood of matching marks from scenes of crime. There are currently 900,000 outstanding crime scene marks on police databases”. This point was repeated by then Prime Minister Tony Blair in a newspaper article (Blair, 2006) suggesting that such secondary uses of the identity database may well have been officially sanctioned.

The government’s proposals for identity cards proved to be unpopular with fewer than 15,000 cards issued by 2010. Indeed, much of the media coverage of the Scheme was very critical of the proposals (Pieri, 2009; Whitley, 2009) and, following the General Election in May 2010 (when only the Labour Party was still supporting the proposals (Whittley & Hosein, 2010c)), the new Coalition Government scrapped the whole scheme, physically destroying the National Identity Register and introducing the Identity Documents Act (2010) that repealed the Identity Cards Act.

The challenge of effective identity policies, for example for accessing government services online, did not disappear with the election of a new government and since May 2010 the UK Cabinet Office has been developing identity assurance (IdA) policies. A report by Sir James Crosby for then Chancellor Gordon Brown provided a useful distinction between identity management and identity assurance whereby ‘identity management’ suggests data sharing and database consolidation, concepts which principally serve the interests of the owner of the database, for example the Government or the banks whereas ‘identity assurance’ is a consumer–led concept, a process that meets an important consumer need without necessarily providing any spin–off benefits to the owner of any database (Sir James Crosby, 2008).

The IdA programme embodies its own political assumptions, not least the realpolitik that any UK identity policy must now align with the policy commitments of supporting privacy and citizen empowerment that formed the basis of the election manifestos of both coalition partners. In particular, it means that the UK identity policy will not be based around a government held register of identities (i.e. the National Identity Register by another name (Maude, 2012)). Instead, the provision of identity assurance services has been moved entirely to the private sector, through commercial identity providers such as the Post Office, banks, mobile phone providers, etc. Government departments then become consumers of these identity services (via a distributed hub that enhances privacy by explicitly breaking the link between the identity provider and the government service). In addition, because the government is keen to develop a thriving marketplace of identity providers, it has told industry that it will not enter the market place itself.

Although the private sector will be providing the identity assurance services, the government provides oversight of the overall process. For example, it has issued a series of “Good practice guides” that detail the requirements for the secure delivery of online public services, for authentication credentials and the means of validating and verifying the identity of an individual in support of online services (Cabinet Office, 2012).

The particular configuration of the UK’s identity assurance programme reveals a series of technological assumptions that are often implicit in other countries’ identity policies, where the state plays a key role as an identity provider. In this configuration, the state is no longer an identity provider—
instead it becomes an attribute verifier for important, state–based attributes such as citizenship, entitlement to social security benefits, etc. As the consumer of identity services that require minimum security levels, the state specifies the requirements that identity providers must meet in order to be qualified to provide those identity services (Cabinet Office, 2012). In so doing, it makes explicit the considerations of ‘trust’ that previously existed (or didn’t exist) between different government departments and the data they provided when “the Government” was itself the identity provider (Lips, 2012; Wilton, 2012).

3. The legal view of technology neutral policies

A recent paper by Koops (2006) provides a useful review of many of the legal arguments put forward for using technology neutral policies “as a starting point”. Koops’ analysis starts with the classic concern that technology specific regulation might rapidly become out of date or obsolete (Bennett Moses, 2011). He also notes that at times of technological “turbulence” legal certainty is a reasonable desire of regulators and industry. Using the example of in–line skates (are skaters ‘pedestrians’ or ‘cyclists’?), he highlights the challenge of defining categories and the need to be able to revise categories as new technologies emerge (Bowker & Star, 1999; Marche, 1991; Whitley et al., 1989).

Koops then provides a series of different categories of what is meant by technology neutral policies. The first of these is whether the policy should be neutral in terms of how it is formulated or in terms of its effect, with most cases focussing on technology neutral effects that might feed back to technology neutral formulations. Reed (2007) describes this as technology indifference and gives the example of copyright law applying whether a copyrighted work was communicated by e–mail or by semaphore flags. There are also clear parallels here with regards to ongoing debates around network neutrality and associated privacy concerns (Cooper, 2011; Ohm, 2010b).

A variation of this concern about the purpose of technology regulation argues that what holds off–line should also hold on–line—paraphrased by Reed (2007) as implementation neutrality (i.e. the regulation should be neutral about whether the process is implemented via digital technology or not). This is perhaps most clearly seen in regulations about the introduction of e–signatures which are explicitly intended to ‘neutralise’ the organisational choice between e–signatures and wet–ink signatures. e–signatures appear frequently in legal discussions of technology neutral policies (Ali, 2009; Koops, 2006; Reed, 2007).

In terms of the consequences of regulation, another variation of technology neutrality (and one which features heavily in debates about network neutrality) is that regulation should not discriminate against certain technologies, but should also not hinder the development of particular technologies. Reed (2007) calls this potential neutrality. In this latter sense, Zittrain’s arguments about generativity (Zittrain, 2008) are particularly relevant.

A third theme in Koops’ classification focuses on technology neutrality as a legislative technique that allows laws to be sufficiently sustainable in order to provide certainty but also explicit about which technologies they are intended to cover (and why) so that whenever there are fundamental changes to the technology it is possible to trigger a revision in the law.

Variations of this approach include the use of ‘enabling’ legislation and secondary legislation / statutory instruments to present and update the details of the legislation as was the case for the UK Identity Cards Act.

4. A technological view of technology neutral policies

Reed (2007), among others, notes that potential neutrality might be undermined by changing business models associated with the use of new technologies. In this context, he discusses the problems faced by e–money initiatives given that the associated EU Directive defines electronic money as value which is “stored on an electronic device” nothing that this fails to differentiate between the payment and credit activities that e–money could facilitate. Nevertheless, he claims that it is possible to “futureproof” regulations so that “these laws are still capable of applying in spite of the changes in technology” (Reed, 2007 p. 276). Arguably this view carries an implicit essentialist assumption that there are particular
attributes of a technology that are “essential to achieve the legal results that the regulator is aiming for” (Reed, 2007 p. 276).

Such essentialist perspectives on technology (Grint & Woolgar, 1997) have been called into question in a range of cases (e.g. Cadili & Whitley, 2005) where, time and again, it is the specifics of the technology that are significant, not generalisable and generalised principles or attributes.

Data protection laws are frequently cited as a classic example of technology neutral policies that are based on such essential principles. National data protection laws in Europe, such as the UK’s Data Protection Act (1998), are local transpositions of the EU Data Protection Directive 95/46/EC. This is itself based on a series of earlier best practice guidelines that were frequently presented in terms of a series of ‘principles’ of data protection.

In the early 1970s “Fair Information Practices” emerged from a report published by the U.S. Department of Health, Education and Welfare (U.S. Department of Health Education and Welfare (HEW), 1973). These principles are technology–independent procedural guarantees which attempt to balance the rights of individuals with those of organizations.

The HEW Fair Information Practices were developed further in a document produced by the OECD in 1980 (OECD, 1980). The document was explicitly designed as a response to the “development of automatic data processing, which enables vast quantities of data to be transmitted within seconds across national frontiers, and indeed across continents” and was presented in terms of “a consensus on basic principles which can be built into existing national legislation, or serve as a basis for legislation in those countries which do not yet have it”.

Within Europe similar considerations gave rise to the introduction of the Council of Europe Convention of 1981 that provided, in turn, the impetus for the first UK Data Protection Act in 1984. A decade later, the EU revised its stance on data protection with a new directive (EU, 1995) and has done so again with proposals to replace the data protection directive with a new regulation (EU, 2012a; EU, 2012b).

The UK’s Information Commissioner’s Office (ICO) presents the UK Data Protection Act in terms of eight Data Protection Principles (Information Commissioner's Office, 2011) and a recent technical report on the Data Protection Directive sponsored by the ICO noted that one of the main strengths of the directive was the fact that it was “flexible due to a principles–based framework” and that it was “technology neutral”, making “no reference to specific technologies” and that its “concept of personal data was broad enough to be technologically neutral” (Robinson et al., 2009 p. 22).

Nevertheless, recent opinions by the EU Article 29 Working Party call into question the ongoing usefulness of the technology neutral principles underlying the directive. For example, online services like social networking sites such as Facebook call into question the distinction between data controller and data processor (Article 29 Data protection working party, 2010). In the mainframe era, when the data protection principles were first formulated, technological considerations meant that the distinction was clear. However, as the Working Party notes:

Social network service providers provide online communication platforms which enable individuals to publish and exchange information with other users. These service providers are data controllers, since they determine both the purposes and the means of the processing of such information. The users of such networks, uploading personal data also of third parties, would qualify as controllers provided that their activities are not subject to the so–called ‘household exception’ (Article 29 Data protection working party, 2010 p. 23).

Other concerns have been raised by the Working Party when exploring what should be considered to be ‘personal data’, particularly in relation to search engines and their retention of IP addresses (Article 29 Data protection working party, 2008), see also (Pounder, 200) and questions of re–identification more generally (Narayanan & Shmatikov, 2010; O’Hara et al., 2011; Ohm, 2010a).

Whilst it is possible to categorise these changes in terms of the evolution of underlying business models, there are also technological developments that produce a much sharper discontinuity from the existing ways of doing things. A number of such discontinuities can be found in the area of e–identity and are discussed more fully in the next section.
5. Technology specific alternatives for e–identity

The argument for general principles underlying technology neutral policies presumes a linear innovation path (or a Kuhnian “normal science”) where technological developments refine and enhance existing technologies. In the case of identity related technologies, this assumption of linearity is no longer valid. Work in cryptographic areas by researchers such as Stefan Brands (Brands, 2000) and others (Auerbach, 2004; Deswarte & Gambis, 2009; Engberg, 2004) offers a Kuhnian paradigm shift in terms of our thinking about identity. As privacy policy expert Caspar Bowden notes, “Since about 1997 we have known that a viable solution is mathematically (really) anonymous credentials, transacted over onion routing, and fleshed out in projects such a PRIME. All these transactions can be done mathematically privately and even audit and performance statistics can be collected (perfectly) privately. Indeed, this is precisely what cryptographic and privacy engineering was invented for, and without it ‘trusted’ parties unnecessarily collect log data which can track a whole nation in a transactional surveillance grid”.

Some, such as privacy and security expert Stephan Engberg (2011), suggest that the use of these cryptographic techniques can transform modern society, enabling individual citizens to reassert meaningful control over their personal data, effecting a sophisticated form of informational self–determination (cf Hornung & Schnabel, 2009). In Engberg’s view, liberal democracy in the digital age “will depend on digital structures that by default are not able to distinguish (end–to–end across purposes) between two transactions by the same person / device and two transactions by different entities”.

Whilst there are practical questions to be addressed in transforming existing identity infrastructures and services to operate in this way (most public services still use “names” as important (secondary) identifiers for the citizens they interact with, not least because that simplifies the customer experience in case of problems with the transaction), arguably a technologically sophisticated country like the UK (and particularly one without legacy centralised identity systems to integrate with) should be looking forward to technologically sophisticated, citizen–centric systems rather than relying on “rear–view mirror” perspectives that, intentionally or not, incorporate technologically unnecessary surveillance capabilities.

Even without moving to the sophistication of systems such as these, there are other forms of technological innovations that do not sit well with the ‘general principles’ of identity management. For example, biometric technologies are widely considered to be essential components of secure, robust and dependable identity assurance, especially in today’s interlinked digital economy where passwords and PINs no longer suffice as a strong means of user authentication.

However, thinking of biometrics in technology neutral terms (“a means of identifying someone”) rather than a means of linking claims to a person has a number of risks, not least in terms of biometric failures.

In her recent book, Shoshana Magnet (2011) reviews the very real ways in which biometric systems can fail. These failures are “failures to meet basic standards of objectivity and neutrality in their application, and the failure to adequately conceive of the human subjects and identities that are their purported objects” (p. 2). That is, “these new identification technologies suffer from ‘demographic failures’, in which they reliably fail to identify particular segments of the population. That is, even though they are sold as able to target markets and sell products to people specifically identified on the basis of their gender and race identities, instead these technologies regularly overtarget, fail to identify and exclude particular communities” (p. 5).

She notes that “biometric technologies that rely on erroneous assumptions about the biological nature of race, gender, and sexuality produce unbiometrifiable bodies, resulting in individuals who are denied their basic human rights to mobility, employment, food and housing” (p. 151).

Another risk associated with an uncritical acceptance of biometrics in technologically neutral terms is that of closing down opportunities for innovation—opportunities which companies such as Touch2id have been exploring.

Touch2id is a private organisation (http://www.touch2id.co.uk) that has been operating its biometric age–verification scheme in Wiltshire, UK since January 2010 in partnership with Wiltshire Council, NHS PCT, Police, Licensing Authority and Trading Standards officials.

Touch2id seeks to use the principles of data minimization and minimal disclosure to use biometric data in a privacy–friendly and effective manner. These principles call for systems that store and reveal the minimum amount of personal data for the transaction at hand. It provides a proof–of–age card, NFC service or mobile phone ‘sticker’ for 18–25 year olds that uses a single fingerprint sensor instead of a picture to confirm the claim that a person is 18 or over. It does not need to store personal details like
name, date–of–birth, gender or address, operates free of a central database and does not capture and store a fingerprint at enrolment or during authentication.

Diversifying from biometric–based access control systems, Touch2id was set up in 2008 to develop a unique, database–free solution for proof–of–age for the UK. The technology harnesses the latest multi–spectral fingerprint sensors which can read the fingerprint from below the epidermis (overcoming typical performance failures in the field due to dirty or damaged fingerprints) and the emerging use of contactless smart–card technology, such as London’s Oyster card and NFC–enabled smart phones. The unique aspect of the technology, which has been developed in the UK with overseas OEM partners, is a database–free application that stores only a unique code created from the fingerprint, using a process known as minutiae mapping. That is, the fingerprint itself is never captured and the approach therefore minimises the risks of unrecoverable biometrics (e.g. Ratha et al., 2007).

Having enrolled with Touch2id (which involves generating the unique fingerprint code and storing it on the credential alongside a verified date–of–birth and other system information), the young person is free to use this credential to prove their age. For example, when entering a bar, they present their credential to a reader device and present their fingerprint on the reader, where a second code is generated from the presented fingerprint. If the codes match and the date–of–birth stored on the chip confirms that the person is over–18 on that day, a green light flashes and sound is generated confirming that the person is over–18.

Touch2id therefore provides a form of zero–knowledge proof for the claim that an individual is over 18. It does so without revealing anything other than the veracity of that claim. The bar owner checking whether someone is over 18 can do so with a sufficient level of assurance in this claim (the date–of–birth has been verified and the person presenting the credential is the person that it was issued to) without having (or needing) to know the individual’s name, address, gender or date–of–birth and without generating an archival record of who visited which bars when.

Both Touch2id and the cryptographic approaches outlined above raise serious questions about the notion of technology neutral e–identity policies. They offer zero–knowledge proofs and are not based on databases or transactional surveillance grids. Whilst, of course, these technologies could be presented using their own technology neutral vocabulary, finding a form of general principles that incorporates both these technologies and the existing identity management techniques found in the marketplace at present would be difficult.

6. From identifiability to levels of assurance

In parallel to the development of innovative technologies, another theme that is being adopted in the UK as elsewhere is an explicit move from the idea of ‘perfect identifiability’ to a risk–based perspective based on required levels of assurance. Just as the bar owner does not need to know the name of the customer only—whether they are over 18—in many cases so–called identification claims are actually authentication claims (Whitley & Hosein, 2010b). That is, a relying party in the interaction must be able to assess the likelihood that the person they are interacting with legitimately has the attributes they claim (Cabinet Office, 2012). As in the case of Touch2id, this attribute might be a claim to be over 18, or being allergic to penicillin or, for online interactions, a ‘real person’ to move past ‘captcha’ screens.

Some of these claims might be self–asserted (such as allergies, or avatar name), others might be backed by a level of assurance provided by a commercial identity provider (a credit reference agency might provide support for claims of credit worthiness) whilst others might require particularly high levels of assurance, for example, claims that a person is a UK citizen and hence entitled to a British passport, or has no criminal record history and is therefore eligible to work with vulnerable people.

Such a risk–based perspective, if followed through logically and taking advantage of recent developments in computational computer science and cryptography, potentially removes the need to use identities for many transactions at all. For example, payments to an online store are currently based around the customer sending their bank details and other personal data to the store (which acts as the relying party for the financial part of the transaction; symmetrically, the customer is also a relying party in terms of the claim that the store will deliver the purchased goods or services). However, logically, all that is required is for the online store to receive an assertion (or guarantee) from the customer’s bank that the bank will cover the claimed payment (once, for that amount at (around) that time / date). If the customer makes another online transaction to a different store, a new one–off guarantee would be issued.
to this new relying party by the bank. Each online store is able to process the payment securely without ever needing to know who the customer was and without the need to receive, and store securely, the customer’s data. A public demonstration of this capability is available at: (Trusted Attribute Aggregation Service Demonstration, 2011).

This perspective, informed by recent developments in technology raise important concerns about arguments for technology neutral policies based on general principles. The next section uses this perspective to review recent proposals on identity policies for the EU.

7. A study on identification, authentication and signature (IAS) policy

Recent work financed by the European Commission studies the feasibility of a comprehensive EU legal framework that would apply to electronic assertions needed to secure electronic transactions as well as the ancillary services needed to use them: electronic identification, authentication, signature, seals and certified delivery. The objective would be to facilitate the smooth working of electronic transactions in the internal market (IAS Project, 2011b).

The draft of the first IAS deliverable echoes the EU’s Digital Agenda (EU, 2010) by explicitly stating that “the policy goals that an IAS approach should cover, including such aspects as the enabling of the internal market, technological neutrality and legal reliability” (IAS Project, 2011a emphasis added).

The study makes many useful recommendations, particularly around the challenges of electronic signatures for the internal market (where the question of being technology neutral between paper and electronic signatures across national boundaries becomes significant). However, there are a number of elements where the attempts to be ‘technology neutral’ from a legal perspective undermine the benefits of the proposals from a technological one.

Given Engberg’s assertion that it should be possible to design systems that are unable to distinguish between two transactions by the same person / device and two transactions by two different entities, many of the claims about uniqueness and identifiability made in the IAS study bear further scrutiny. For example, when discussing Electronic Identity Establishment the report argues that enrolment “relates those (identity) attributes to a primary key (electronic identity primary key, i.e. an identifier consisting in itself a Unique Identity) for later retrieval”. It continues: “There is typically a repository or database that may be centralised or decentralised in nature” (IAS Project, 2011a p. 23 emphasis added). As noted above, no such repository or database should be needed, nor should it be necessary (or even desirable) to try to identify or define a ‘unique identity’.

This focus on unique identification also affects the description of biometrics, where the report presents the wish that: “in the ideal world there would be a technology providing some inimitable unique identity (IUI) to every natural person”. This IUI would have the properties of being a “Unique Identity of the entity it is related to; 100% unique to that entity; derived from biometric properties with 100% reliability for a lifetime; be as short as possible” (IAS Project, 2011a p. 25). Empirical studies of biometrics have shown how dependent they are on all sorts of contextual factors, including the technology used to capture the biometric image (Bower et al., 2009; Magnet, 2011). In addition there are important conceptual questions about what uniqueness means in any particular context (Cole, 2009). Moreover, as the UK Identity Assurance Programme shows, it is possible (through the use of required levels of assurance) to allow individuals to use multiple credentials, from multiple identity assurance providers with different levels of assurance rather than requiring the system to fixate on some inimitable unique identity that is used for all identity related transactions regardless of the required level of assurance.

The report states that “the current state–of–the–art can certainly rely on the concept of Unique Identity … while trying to evolve towards Inimitable Unique Identity in the future”. The intention is to link this unique identity, through a one–way function, to a Unique Identity Derivation “a special type of electronic identity primary key as named above, i.e. a special unique identifier consisting in itself a Unique Identity of the entity it is related to” (IAS Project, 2011a p. 25).

Moreover, even this disguised identity derivation is, according to the study, intended to be used by Identity Attribute Assertion Providers who assign identity attributes to persons in a way that would “ideally be 100% in the context of an official Governmental identification scheme but may vary between different levels in more relaxed market or business or social application domains” (IAS Project, 2011a p. 26).
Achieving close to 100% certainty in a unique identity is always a costly process, particularly because of the opportunities for fraud that are opened up if an unique identity is incorrectly assigned. Moreover, in addition to the removal of a technological need for 100% identifiability, a risk–based perspective also reduces the emphasis on 100% certainty, instead any service provider (whether public or private sector) needs to assess the level of assurance it requires from an identity provider and these will clearly vary from context to context (Cabinet Office, 2012).

For example, the report gives a number of use cases where entity authentication is seen as “a process of establishing an acceptable level of assurance that a claimed identity is genuine” (IAS Project, 2011a p. 28). However, many of the use cases do not actually require identity, only support for the claims being made. That is, many are transactions that could easily be achieved through a form of pseudonymity rather than identity. These are transactions where the relying party does not need to know ‘who’ they are dealing with, only that they have the required attributes—see Error! Reference source not found. for an analysis of these use cases.

<table>
<thead>
<tr>
<th>Table 1 Identification Use Cases (IAS Project 2011a, pp. 28-29)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Use Case</strong></td>
</tr>
<tr>
<td>Checking e-mail and voice mail</td>
</tr>
<tr>
<td>Voting</td>
</tr>
<tr>
<td>Border control</td>
</tr>
<tr>
<td>Listening to streaming music, through a pay subscription</td>
</tr>
<tr>
<td>Internet setting (including urgent access to emergency services)</td>
</tr>
<tr>
<td>Managing the domestic services of your home</td>
</tr>
<tr>
<td>Log on to an electronic service provider or Government service for being eligible for benefits (address, family status, age etc.)</td>
</tr>
</tbody>
</table>

This example in fact confuse the possession of a credential which supports the claim (which is all the benefits agency requires) with the information used to issue the credential (address, family status, age etc.).
In many of the use cases, pseudonymous transactions based on attributes are feasible and possibly desirable. These transactions are not truly anonymous, rather they are cases where the pseudonymity of the transaction is maintained by the identity provider that is supporting the claims. That is, if my bank knows that I am credit worthy then it should be able to support my claim to make be able to make purchases online (for example by issuing me with a credit card for online purchases), even if the relying party sees my (pseudonymous) name (“Mickey Mouse”) rather than my real name. If a transaction is believed to be fraudulent, it may become necessary to break the pseudonymisation and the identity provider will be able to do this, on presentation of a suitably authorised ‘search’ warrant.

Another example given in the study is where social networks or other service providers are used to support attribute assertions and the example of a mobile phone number and a phone bill is used to identify and authenticate a person applying for a bank loan. Once again, this example confuses two things, identifying the person and supporting their suitability for credit. Whilst mobile phone SIM cards are regulated in some countries (i.e. they can only be sold on presentation of official identification documents) in others there is no such restriction. Indeed, it is possible to buy SIM cards in UK airports (both airside and terminal side) from automatic dispensers, for cash.

The discussion of e–signature types again reveals some technological assumptions, which originate in the previous e–signatures directive (EU, 1999). In the original directive, an ‘advanced electronic signature’ means an electronic signature which satisfies various requirements, such as being uniquely linked to the signatory and is created using means that the signatory can maintain under his sole control. However it also states that an ‘advanced electronic signature’ ”(b) is capable of identifying the signatory”.

As the discussion of technological developments presented above has shown, it is now possible to implement functionality which addresses all the requirements of advanced electronic signatures and more, but which either are capable of pseudonymously identifying the signatory or which do not need to identify the signatory at all to maintain their utility. As such, simply restating this particular requirement highlights the technological dangers of ‘technology neutral’ policies.

8. Concluding discussion

The examples presented in the paper suggest that the unqualified support for technology neutral policy for e–identity needs to be reconsidered. Developments in technology as well as innovative business models that are explicitly designed to be privacy–friendly have the potential to reshape the very form of e–identity policy. As such, the ‘general principles’ underlying e–identity have been subject to a paradigm shift and it is difficult to reconcile the new principles of e–identity with those found in earlier technologies. Whilst it is possible to formulate new general principles, a more straightforward response might be to recognise that general principles or technology neutral policies should be restricted to being a starting point rather than a desirable end point. That is, in line with Koops’ potential neutrality technology specific regulations should be reviewed regularly to ensure that they are not unnecessarily distorting the market and innovation opportunities by supporting or restricting particular developments (Bennett Moses, 2011).

It also suggests that technology neutrality is not a simple binary choice—these regulations are technology neutral / those are not. Instead, technology neutrality might end up being more a question of degree.

A related concern with technologically neutral policy in practice is that it shifts detailed, technology–specific decisions to secondary legislation, statutory instruments and their equivalents. Although, in principle, such mechanisms are entirely appropriate, in practice they face far less parliamentary scrutiny even though they may have far greater practical implications. This unintended consequence of the neutrality argument, where detailed considerations are made by unnamed technocrats and civil servants rather than the elected legislature, is inherently problematic and open to abuse. Similar problems arise when these issues are transferred to international standards bodies that claim to act on behalf of citizens.

Technology neutral policies also raise significant concerns for techno–legal integration across EU member states. The certainty about legal and technical interoperability can only be achieved by being specific about technology and regulations. Failing to do so might achieve political agreement but runs the risk of low levels of take–up and adoption.
By removing the state from being an identity provider, the UK experience brings to the forefront questions of liability that might remain implicit if the state is intimately involved in both the issuing and use of identity credentials.

The technology alternatives presented in this paper are particularly amenable to user-centric considerations. Rather than being restricted to a single identity provider (the state), UK citizens are intended to be able to take more control over their identity credentials, organising them in ways which support their own needs rather than those of the state. For example, a citizen may choose to use some one credential for work-related activities, another for family and household identity claims, a third for their hobbies etc. Moreover, the technologies discussed above allow for various forms of anonymous (e.g. Touch2id) and pseudonymous actions. These will have knock-on effects on legal models (and liability), particularly in the case of zero-knowledge proofs. What underlying law should apply if a zero-knowledge proof is used? Who is liable in cases of dispute?

Koops ends his review by asking four questions when thinking of using technology neutral regulation “as a starting point”: What is the goal of the regulation? Is it desirable to control technology? What level of legal certainty is required? And how urgent is the need for regulation?

In the context of e-identity policies it is clear that the goal of regulation should not be to provide equivalence between online and offline worlds. Indeed, as was shown in the paper, recent technological developments offer the opportunity for enhancing privacy whilst supporting identity claims in ways that have no meaningful equivalence in offline worlds.

The role of a market of identity providers in the UK programme highlights the tension between the need to control, at one level, both the companies and the services they provide, while explicitly not restricting new, innovative companies from entering the market place. The associated level of legal certainty is therefore provided by the accreditation schemes associated with becoming an authorised identity provider that can provide identity services to public service providers.

What each of these examples has shown is that in addition to legal concerns about the applicability of technology neutral policies there are increasingly important technological concerns that limit this applicability. If e-identity, and similar technologically sophisticated services, are to succeed we need to reflect critically on the assumption that technology neutral policies are the most effective way forward.

References


On technology neutral policies for e–identity: A critical reflection based on UK identity policy


EU (2010) Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A Digital Agenda for Europe.


Trusted Attribute Aggregation Service Demonstration (2011) Archived at http://sec.cs.kent.ac.uk/demos


On technology neutral policies for e–identity: A critical reflection based on UK identity policy


* * * * *

© 2013 This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works.