Information Commissioner's Office

Promoting public access to official information
and protecting your personal information

New approaches to identity management and privacy

A guide prepared for the Information Commissioner

December 2007

Prepared by:

John Harrison, Edentity Ltd
Pete Bramhall, HP Labs

Table of contents

Acknowledgements

The authors thank Robin Wilton of Sun Microsystems, and Jerry Fishenden of Microsoft, for their contributions to this paper.

Preface

The collection and storage of large amounts of personal information in databases, as well as the possible sharing of that information amongst different organisations, leads to increased risks for individuals. The more information that is held and/or shared, the greater the potential impact on an individual were that information to be lost or stolen.

Identity management – the use of personal information in order to adequately identify an individual who is trying to access your products or services – is an important tool as it not only protects organisations through ensuring that the right person is accessing the right services but can also be a way of helping to protect individuals from the risks that large databases and increased information sharing bring through giving some control and choice back to the individual.

Traditionally, identity management has centred on allocating a single identifier to an individual and putting all the information about them in a database. However, there are other identity management techniques which can reduce the risk posed by putting information all together in one collection or using a single, common identifier.

This paper was commissioned by the ICO from experts with experience of the different identity management techniques in order to describe the different solutions available which can bring protection to individuals and the situations in which these solutions can be appropriately applied. Different identity management techniques will be applicable in different situations. The purpose of this paper is to help inform those working in this area of the alternative forms of identity management techniques that are available and which can help to protect privacy.

Whilst the paper was commissioned by the ICO, it is aimed at broadening the understanding of the identity management techniques available rather than the Commissioner advocating a particular technique in all circumstances.

Introduction & Summary

1. The Information Commissioner is responsible, among other activities, for the UK's data protection regime. The role's origins date back to 1984 when organisations were fast learning how to use computers to keep records about their customers, and the cost of sharing data was falling dramatically. To guard against the risk to privacy, Parliament passed the original Data Protection Act and so created the predecessor of the Information Commissioner's Office, the Data Protection Registrar.

2. Then, as now, consent mattered. The idea that an organisation can legitimise a proposed use of personal data by obtaining explicit consent from the individual to whom it relates is implicit in the principles of data protection. And there are many examples of consent in action. Consider:

   - the tick-box at the end of an application form, whereby an individual can indicate whether or not she wishes to receive marketing information from 'selected partner organisations';
   - that most banks require, by means of a clause in the application forms for new accounts, that customers agree to the sharing of data with other banks. Generally, they use the credit reference agencies as intermediaries for this data sharing.

   It can be argued that 'consent' is too general a word to cover both these examples. In the first, the individual has a real choice; in the second he is faced with a Hobson's choice: either agree to data sharing or go without a bank account. But this is a fine distinction, at least for these purposes.

3. What is important, in both cases, is that the individual uses what might be called a '*front-office'* communication channel – either the post, telephone, internet, or face-to-face – to give the 'consent', and then the organisations use different '*back-office'* channels for actually sharing the data. This separation of the channels – front-office for consent, and back-office for the actual sharing of data – limits the immediacy and granularity with which consent can be given, requires that the individual place a high degree of trust in the organisation to which consent is given, and can present the organisations sharing data with significant record matching problems.

4. But technology moves on. Back in the 1980s, personal computers were rare, and the internet was only used by computer scientists. Now, according to the latest Oxford Internet Survey [OII] sixty-six percent of Britons have access to the internet at home, and – of these eighty-five percent use a broadband connection, rather than the slower dial-up technology. This rapid adoption of new technology offers society the scope for further change, towards a world in which individuals can act as gate-keepers for the use of their own personal data.

5. In this world, permission for data-sharing, and the data itself, will pass over the same front-office channels. Note the terminology here: we use the word '*permission'*, in lieu of '*consent'*, to imply that the individual is 'in the loop', and has fine-grained, active, and transaction-specific control over the sharing of her data. Other terms, such as '*user-centric'*, or '*citizen-mediated'*, or '*front-office*' data-sharing also have some currency.

6. In one sense, permissioned data sharing is nothing new. Although negative information – such as criminal records – must always be shared directly between organisations, individuals have long managed certain kinds of personal information themselves, using tamper-proof paper certificates. Examples include medical prescriptions, driving licences, and exam certificates. Generally, such information is positive, or can at least be used for transactions that deliver a net benefit for the individual.

7. Replicating the permissioned approach electronically requires an emerging technology called '*user-centric identity management*'. Because of confusion with the better established, but distinct, technology of '*organisation-centric identity management',* user-centric IdM is poorly understood. Other confounding factors include the inevitable uncertainties that arise in new fields of work, and the fog created by competing sources of information, all using terminology in subtly different ways. As one example, the term 'federated identity management' is common in the field, but now means so many different things to different people that it has ceased to be useful.

8. This guide, written for the Information Commissioner, starts by explaining how the two approaches to identity management work, how they differ, and where each can be used. Next there is a description of two early implementations of the user-centric approach at the national scale, followed by a survey of competing initiatives to create true user-centric IdM infrastructure. The document concludes with the recognition that such infrastructure is unlikely to reach maturity without active involvement from the public sector.

From data sharing to identity management

9. "*It's simple. Everyone has an identity, don't they? My identity is my first name and surname, and that's how everyone I know knows me.*" Well, not quite. A name, such as John Smith, is not really an identity. It's an identifier, just one part of an identity. It can be used by John Smith to distinguish himself in a fairly small group. But as the group gets larger, there are likely to be several people sharing the same name, and confusion will follow. To fix the problem, it's necessary to use identifiers that are guaranteed to be unique within the group. These are often called '*customer reference numbers'*, '*identity numbers'* or '*usernames'*, but we will just call them '*unique identifiers'*.

10. Every one of us deals with many distinct entities. Some are in the public-sector; some are businesses; some are charities; some are even distinct parts of, or systems within, a single organisation. Since most entities issue a unique identifier for their own purposes, every individual will find herself labelled with many different identifiers. This cuts both ways: it's good because it protects privacy by making it harder for unscrupulous entities to share data about an individual without consent; but it's bad if entities that should be able to share data, for the benefit either of society or the individual, cannot do so.

Organisation-centric IdM

11. Once this fork is recognised, it becomes clear that there are two routes forward. The first – '*organisation-centric (or enterprise) identity management'* – tends to be adopted by entities that are closely linked in one way or another. They might be different divisions or subsidiaries within a large commercial company, different

public sector organisations involved in a joint initiative, or different parts of a local authority.

12. The first step is to draw (perhaps only as a mental picture) a new organisational boundary around the group as a whole. Then it's necessary to find a way to match-up all the records about an individual. This can be difficult, because the records may be held in different databases, be stored in different formats, and use inconsistent data: out-of-date addresses, and different name formats – Jim or James? – are common examples of this last problem. These days, record matching is usually carried out by computer, using probabilistic pattern-matching algorithms, and human intervention only where necessary.

13. Once the records have been matched, the individual will be likely be tagged with a new identifier for use by all entities within the group. And the individual, if she interacts with the group online, may be given be a single online account, through which to inspect all the data held by the group. In effect, the group now has created a single record about that person, possibly distributed over multiple databases, but unified logically by the common identifier.

14. As the scope of the single record increases, entities within a group may use internal authorisation procedures to provide some safeguards for privacy. This is particularly common in the public sector, where access to a citizen's record is generally limited to those civil servants who have a legitimate need. In some cases, groups go further, and use detailed authorisation procedures to limit access to particular *parts* of an individual's record. The effectiveness of such procedures in protecting privacy tends to fall as the number of employees (and possibly others) who are authorised to view a record increases, and also as the incentives for gaining unauthorised access rise. Moreover, the acceptability of the approach to the public depends on the reputation of the organisation for acting ethically.

15. In this kind of discussion, it can be useful to replace the loaded word '*identity*' with the less controversial word '*relationship*'. Few people argue with the idea that it is sensible for an individual to have a single relationship with an organisation, and that the organisation should be able to enjoy a single view of that relationship, rather than have bits of it scattered around different systems. Seen this way, organisation-centric identity management is a close relative of a well-established, and uncontroversial, business activity, 'c*ustomer relationship management*' *(CRM).* But controversy does arise when – for reasons of efficiency – more and more entities join together for CRM purposes, and all have sight of an ever-expanding single view of the customer. Similarly, some people are concerned when an organisation, with which they have a relationship, contracts with other organisations to provide outsourced business services.

User-centric IdM

16. Despite the trend towards data sharing and record-centralisation, there will always – at least in Western democracies – be a multiplicity of entities which, for privacy reasons, should each retain a separate relationship with the individual, and not attempt to share data or create aggregate records without explicit consent. This brings us to the second IdM approach, '*user-centric identity management'.* This is sometimes also referred to as *'Identity 2.0'*, and is neatly summarised in an entertaining online presentation by Dick Hardt [HARDT].

17. The emphasis in the user-centric model is that the individual should continue to have multiple identifiers, one for each distinct entity with which he or she transacts. Thus, for example, the identifier carried on a national ID card becomes just one of many of his or her identifiers, although it may often be referred to as an individual's '*legal identity'.* The information associated with each identifier is often referred to as a '*partial identity'* or a '*persona'*. The absence within these partial identities of any common unique identifier makes it more difficult for the entities to share data without permission, thus protecting privacy and addressing concerns [ICO] about the emergence of a surveillance society. But it leaves unresolved the issues of (i) how to provide the individual with the convenience of secure single sign-on to multiple distinct entities; and (ii) how to enable the individual to give – in an active way – fine-grained permission for the sharing of specific personal data between such entities when it is to their advantage to do so.

18. Fortunately, the solutions to these two issues complement each other. To enable single-sign-on, it's necessary to separate the processes of authentication and identification. Authentication becomes a quasi-anonymous process by which an individual authenticates to, and thereby demonstrates rights to the use of, what can be called a '*point of network presence'.* The individual's relationships can be thought of as lines radiating from the point of presence to the various counterparties, each line distinct and tagged with the identifier used by a particular counterparty. Once signed into her point of presence, an individual can click through to any desired counterparty, provided always that her level of authentication is acceptable to the counterparty.

19. This same point of network presence – which can be held on either a personal device or on a server – can be used to provide the individual with a '*permission hub'*, with which to control precisely how much personal information is revealed to any counterparty. But, as a necessary corollary, the counterparty can always decline to complete a transaction if the information revealed is insufficient. A few examples might help convey the idea:

    - to hire a car, an individual must show his driving licence to a rental company. Using a permission hub, linked to the driving licence registry, he could prove to the rental company online that he is licensed to drive, and so reduce the paperwork necessary when picking up the car.
    - to open a bank account, an individual must show the bank proof of his legal identity, typically a passport or birth certificate. Using a permission hub, linked to a register of legal identities held by government, he could carry out the transaction entirely online, avoiding the need to mail documents or visit a bank in person.
    - in some countries, anyone applying for a new job must show proof of claimed qualifications. Using a permission hub, linked to the databases operated by the awarding bodies, it would again be possible to carry out this transaction online, and eliminate the use of paper.
    - to purchase alcohol online in many countries, an individual must prove that he is above a certain minimum age. Using a permission-hub, he could do this, without necessarily disclosing his precise age or date-of-birth. Similarly, an individual could demonstrate that he is young enough to enter a teenage chat-room.

- finally, an individual could use a permission-hub to demonstrate that he is entitled to free healthcare, perhaps of a sensitive nature, without revealing his legal identity; and could prove that he is entitled to vote in an election, and then do so anonymously.

20. Note, in particular, that a permission-hub empowers the individual to control precisely how much identifying information is revealed in any given transaction. The choices for an individual are:
    - to be completely anonymous, using a different identifier (or no identifier at all) on each occasion that he visits a given counterparty;
    - to be pseudonymous, using the same identifier every time he visits a given counterparty, but different identifiers for different counterparties; and
    - to use a common identifier for multiple counterparties, or permit a cross-organisational identifier to be transmitted between counterparties through his permission-hub (which amounts to the same thing.) By doing so, he would be inviting such counterparties to merge their records about him. This is the online equivalent of showing one organisation an 'identity card' given to you by another.
21. There are many further examples of transactions that can be enabled by user-centric IdM. Some are direct analogues of transactions that already take place using paper, while others are only made possible by the speed of electronic networks. No one application can be developed in isolation of all the others: what is required is the creation of new infrastructure on which the many different applications can be built.
22. In the long run, the cost to the economy of not creating this new infrastructure may be significant. Recent research [TRGD] has found that some individuals will choose not to conduct a transaction online if they if they fear that they will lose control over the data that they are required to submit. A user-centric IdM infrastructure offers that control.
23. In later sections, two early applications of the user-centric approach in national-scale systems are described, and are followed by a progress report on initiatives to create true user-centric infrastructure. But first it's useful to emphasise the ways in which organisation-centric and user-centric approaches to IdM differ.

Difference One: Authentication and identification

24. The two approaches to IdM place different emphases on the processes of '*identification'* and '*authentication*'. Because back-office data-sharing, and organisation-centric IdM, require the receiving entity to match incoming files, or a person in the flesh, against records already held, it's necessary to devise a system that allows all the records, and the individual, to be linked together. As described above, a common solution is to create a system of unique identifiers—such as the NHS number, the proposed Unique Learner Number, or the National ID Card number itself – that appears on every record, and thus makes the matching process easy. Matching need not involve the individual at all, and typically does not when the data is negative and the individual thus has no discretion over whether it should be shared: criminal records are the obvious example here.

25. But when the data is positive, or at least neutral, it can be useful to give the individual a card, bearing the relevant number, which he can present as a means of linking himself to his aggregate record. Individuals tend to accept, or even welcome, such cards provided that they are willing to maintain a single relationship with the grouping of organisations that has got together for the purposes of the card scheme. In other words, the individual should know in advance which organisations are going to make use of any given card scheme, and be content that these organisations should share data about him between their back-office systems.

26. Note that cards serve two purposes:

    - as a means of identification, i.e. they convey an identifier – normally a number – that uniquely distinguishes between the bearer and other individuals within the relevant population; and
    - as a means of authentication, i.e. they carry features – such as a photograph, a signature, or even a fingerprint encoded on an integrated chip – which allow the individual to prove that he is the rightful owner of the card.

27. The two processes, identification and authentication, need not always be so closely bound together. Indeed, it's clearly possible to identify an individual without inviting him to authenticate himself. And, less obviously, it's also possible to authenticate someone without identifying them. As an example, suppose that – before the enactment of the present 'Know Your Customer' regulations – an individual approached a bank with a paper cheque, made out to the (anonymous) bearer for a large sum of cash. How can the bank be sure that the bearer is the rightful owner of the cheque, and hasn't stolen it from someone else? Anticipating this problem, the issuer could have implemented various authentication measures, such as affixing one or more of the rightful owner's biometrics – such as an (illegible) signature, a photograph, or a fingerprint – to the cheque, and requiring the bank to verify these before handing over the cash. Note that the bank does not match the bearer's biometrics against some pre-existing database: they are used for authentication, i.e. to prove ownership, rather than for identification.

28. User-centric IdM can be thought of as a way of capitalising on this separability of authentication and identification. Authentication becomes the anonymous process by which an individual proves ownership of his point-of-presence/ permission-hub; and – because the individual 'carries' her own records between organisations – there is no record-matching problem, and no need for any system of cross-organisational identifiers. Instead each organisation need only be concerned with its own internal identifiers, each referring to a unique *relationship* between the customer's point of presence and its own systems.

29. Suppose now that the bearer cheque transaction, described above, was carried out electronically over a user-centric IdM infrastructure. Similar techniques could be used for authentication. But both the encashing bank, and the issuer of the bearer cheque, might well be unwilling to bear the cost of dealing with electronic biometrics, and could instead choose to outsource the job to a third party who is trusted by both. And if this third party is also the organisation used by the individual to provide a single-sign-on service and a permission hub, then the acts of proving ownership of the electronic cheque, and of accessing the single-sign-on service / permission hub, become one and the same.

30. To summarise, the point is that, in user-centric data sharing and identity management, it is the process of proving ownership – authentication – that is important; and that cross-organisational identification is not required. Further, it's necessary to make authentication more secure as either (i) the value /sensitivity of data shared over user-centric infrastructure increases; or (ii) the number of relationships managed by the individual from a single point of presence increases.

31. Provided that an appropriate business model is put in place for user-centric infrastructure, the cost of secure authentication can easily be shared among the many parties who benefit from it. In contrast, an organisation than attempts to bind secure authentication and identification together in one card is likely to find it difficult to persuade other organisations to share the associated costs.

32. That said, it is possible, in principle, for a 'smart' identification card to also be used to provide the secure authentication required by user-centric data sharing. Each individual would be issued with a small card reader that enables a smart card to function as a one-time password (OTP) generator, much like the key-fob devices issued by some banks for secure access to online banking. The process of authentication enabled by a card is anonymous, confirming only that the holder is the rightful owner of a permission-hub. The fact that the card can also serve as an offline identification token is relevant only to the extent that it may be difficult to persuade individuals that something which – potentially – harms privacy offline could contribute to the enhancement of privacy online.

33. As mentioned above, offline identification can be carried out without the involvement of the individual; whereas online authentication requires the individual to participate actively. This difference is key to determining for which applications the two approaches to IdM and data sharing can best be used.

Difference Two: Discretionary and non-discretionary data sharing

34. Neither approach to IdM and data sharing is sufficient by itself. Individuals can only be relied upon to transmit data that shows them in a good light; and yet for the good of society, it's necessary that certain organisations share personal data which is not positive, or even neutral, but negative. To take two examples, criminal records and child-protection information will continue to rely upon a non-discretionary, organisation-centric, approach to IdM and data sharing.

35. Note, however, that the absence of negative information can be treated as positive information. For example, it may become possible to transform the systems by which society ensures that only appropriate individuals get to work with children and the vulnerable: instead of carrying out a 'black-list' check every time a worker changes jobs, why not create a 'white-list' of accredited workers, and enable each worker to disclose his white-list accreditation (and any subsequent revocation) online to potential employers ?

36. All told, a user-centric approach to IdM can be used wherever the required personal information about an individual either positive or neutral, or where – according to the custom of a particular society – the individual is allowed discretion as to whether to disclose information or not. For example, the user-centric approach could be used:

- in education, where most qualifications are seen as positive, and where – in the UK at least – it is accepted that an individual need not disclose bad exam results
- in some areas of the health sector, where there are advantages in having individuals maintain their own records and allow new healthcare providers access to them; and
- in commerce, where the individual would be able – without disclosing any permanent identifiers or communication addresses – to request merchants to send marketing information about a particular kind of product, but only during the precise window of time when they were interested

All of these applications – and the many others that have not been mentioned – rely upon the fundamental user-centric application, secure single-sign-on.

Difference Three: User-centric plays catch-up

37. In the paper-based world, user-centric data sharing was easier and cheaper than the alternative, and was used whenever possible. Thus an individual carried medical prescriptions from doctor to pharmacist, exam certificates from school to sceptical employer, and a driving licence from the DVLA to a car-hire company. Back-office sharing was limited to attributes that were regarded as confidential, such as some medical records; or were negative, such as criminal records.

38. In the network world, the technologies for back-office data sharing and organisation-centric IdM have been developed steadily over the last forty years, as organisations started to use computers and data networks. In contrast, individuals have only acquired network access in the last few years, and so techniques for user-centric data sharing and IdM are only now beginning to catch-up. Also, it turns out that – because user-centric IdM requires infrastructure – it can only be fully be developed with the cooperation of the public sector. This contrasts sharply with the development path of most ICT technologies, which are first proven in the private sector.

39. There's also the point that different environments require different solutions. The use of a single identifier for an individual across multiple organisations appears, to many individuals, to be common sense, and indeed it was appropriate when communication was either face-to-face, or by slow paper-based means. But the network world is different: computers, and high-bandwidth networks, give malicious entities a far greater ability to correlate data from many different sources. In this new environment, it is often argued that individuals need better tools to protect their privacy and manage their own data.

40. The next section starts by looking at two early applications of the user-centric approach to IdM, and is followed by an overview of the various user-centric IdM initiatives that are under development.

Early applications of the user-centric approach at the national level

41. One of the best known, if limited, implementations of user-centric IdM for government purposes is found in Austria. There the government has issued a 'Citizen Card' as part of an electronic identification scheme.

42. Austria does not require a national identity card, but does maintain a single national register of citizens, and an associated system of unique identifiers. The

Citizen Card is best understood as an optional enabler of electronic services. Various steps are taken to avoid the privacy problems associated with a single electronic identifier:

- the (private) identifier stored on the card is generated by adding a seed value to the citizen's public identifier, as recorded on the national register, and then encrypting the result.
- each organisation, or grouping of organisations, in the public administration is allocated a unique 'sector code'. When a citizen uses his card to transact with a public organisation, the sector code is combined with the private identifier held on the card, and the result is then 'hashed' to create a unique sector-specific-identifier.

43. This approach provides strong protection of privacy, since a citizen's sector-specific identifiers cannot be linked across different e-Government services, either directly or by attempting to recalculate the source identifier held on the card. Moreover, the fact that a citizen possesses a card, and so can generate a sector-specific identifier, will indicate to a service provider that the individual is legally entitled to certain privileges associated with an Austrian citizenship or residence, without the need for such service providers to record a single unique national identifier.

44. In the UK, a service called Government Gateway provides the citizen with some protection for privacy. Gateway was commissioned in 2001, as a way of forestalling efforts by individual government departments to create in-house IdM solutions, and so showering citizens with many different authentication credentials. Instead, a citizen can use a single set of Gateway credentials to gain access to services offered by a variety of departments. Most opt for username and password, although a few pay for the greater security of a PKI digital certificate. Like the Austrian e-id scheme, Gateway employs organisation/sector specific identifiers, and so hinders record linkage across different e-government services. Gateway is now being rolled out to local authorities, as part of the broader Government Connects programme.

45. Both Gateway and the Austrian e-id scheme provide a privacy-enhancing SSO service. What they do not do, albeit for different reasons, is provide a permission-hub to enable individuals to act as gatekeeper for control over transmission of their own data between service providers. It can be argued that Gateway is incorrectly positioned for this role: it sits on Government servers, and so – in the eyes of some – cannot serve as a trusted guardian of their data and relationships. And the Austrian e-id scheme uses a card as the point of network presence and so, at first sight, lacks the capacity to offer a permission-hub interface.

User-centric IdM initiatives

46. Over the last several years, various companies, research groups, and open-source enthusiasts have pressed ahead with initiatives to develop user-centric technologies and infrastructures for IdM, as described below:

Information Cards, and Microsoft's CardSpace™

47. Following the failure of the company's initial effort in the IdM space, Passport, and the cancellation of its successor Hailstorm before launch, Microsoft went

quiet about digital identity. Then, in 2004, one Kim Cameron emerged as the company's architect for identity and access and provoked a flurry of discussion in the identity community by blogging about seven 'laws' [CAM], to which – he maintained – any emerging identity system must conform.

48. In early 2005, Cameron began to speak about Microsoft's next identity initiative, an 'identity meta-system' coupled with an interface, called generically 'Information Cards' or 'InfoCards', for use on a personal client device (e.g., a desktop or laptop PC). Microsoft's own implementation of Information Cards, called 'Cardspace™' [CS] is bundled with both the current release of Windows for PCs, Vista, and with Internet Explorer 7; there is also a download available for Windows XP.

49. As its name suggests, InfoCards is built around a card metaphor, and invites an individual to select which 'card' of attributes she wishes to disclose in which contexts. Attributes can either be asserted by the individual, or by a third party. In the latter case, the InfoCards software on the personal device holds pointers to the attributes on the third party's servers, rather than the attributes themselves.

50. The design of InfoCards is inherently privacy-friendly, and can be extended to support applications where attributes are conveyed anonymously, without requiring any identifying information. Also, InfoCards offers protection against the common security attacks of phishing and pharming, and enables the individual to move away from the multiple username/password problem.

51. InfoCards is, clearly, a user-centric approach to identity management. Microsoft learned from experience that service providers will not allow it to be a mass-market server-side intermediary, and so focussed its efforts on a client-side approach, building on the high market share of the Windows operating system. But the decision to take this approach must have been problematic, for several reasons:

    - using InfoCards, an individual's computer itself becomes a factor required for secure authentication, despite the fact that personal computers are insecure, and so are unlikely to be trusted by some counterparties as a sole authentication source;
    - the meta-data within InfoCards will need to be backed-up, and synchronised across multiple devices, which tasks are frequently overlooked by individuals;
    - there are applications for which it is convenient for an individual's permission-hub to be permanently online; and
    - InfoCards, by itself, does not address the needs for a clear server-side business model, or an inter-organisational trust model, and thus does not offer an easy adoption route for organisations.

52. Despite these shortcomings, Infocards should not be discounted. Microsoft is working to make the system usable on mobile 'phones and with smart cards, and to address the back-up issue. Further, the company has stated that it sees as InfoCards as an industry-wide initiative, based on open specifications, from which it will not profit directly. Other companies, such as Apple and Novell, have already released their own implementations. These advantages may persuade public- and private-sector organisations to trial the system. Nonetheless, take-up appears to be slow, and an another approach, OpenID, has attracted significant attention.

OpenID

53. OpenID is a decentralised framework for user-centric IdM, instigated by Brad Fitzpatrick of SixApart, which claims to have become the de-facto standard for the internet. Many experts dispute this claim. The standards are maintained by an emerging OpenID Foundation [OID]. Usage figures are difficult to estimate but, in April 07, myopenid.com estimated that some 2,500 sites were OpenID enabled.

54. The principles behind the current version of OpenID are simple. An individual creates an OpenID account with a provider (A), and then – when asked to register at a new OpenID-enabled site (B) – he gives B a pointer (in the form of a web address) to his account at A. His browser is then redirected to A which checks that he does indeed want to release information to B, and does the deed. Note that, because OpenID requires the presence of an Identity Provider (A) in the transaction, the individual's privacy is only protected to the extent that he trusts A. CardSpace escapes this problem of reliance on a third party, since the permission hub resides on an individual's personal device.

55. OpenID does two things well: it provides a distributed registration and single-sign-on utility for the web. As such, it is lightweight and well-suited for the blogging community from which it sprang, and which it mainly serves. However, as the community considers adding additional functionality for OpenID 2.0, they are beginning to encounter problems familiar to those who have been active in the field for some years. Specifically, they need to (i) overcome the shortcomings in privacy, trust and security that have been pointed by various critics [e.g. ITW]; (ii) consider whether the avoidance of a governance structure, a business model, and a liability model, are sustainable; and (iii) sort out further technical standards for permissioned two-way attribute exchange, for the use of privacy-enhancing relationship identifiers, for reverse marketing, etc.

56. Microsoft has announced support for OpenID; and Kim Cameron has described how a joint usage of CardSpace and OpenID can help overcome some of the latter's problems. However, it remains uncertain whether the scheme can grow beyond the blogging community, and be used for more weighty applications.

The Liberty Alliance and OASIS SAML

57. In 2001, at about the time that Microsoft was trying to win support for its unsuccessful Hailstorm initiative, a number of technology and consumer-facing companies came together in order to establish 'an open-standard for federated network identity through open technical specifications'. The grouping is called the Liberty Alliance [LA] and now counts Sun Microsystems, HP, Intel, Nokia, Ericsson, Amex, NTT, Vodafone and a number of public bodies among its members.

58. Liberty chose to divide its task into three main areas: a federation framework, a web-services framework, and service interface specifications. Work in the first of these areas has already been donated to a standardisation organisation, OASIS, and will inform the development of the SAML standards [SAML]. It seems likely that standards in the remaining two areas will follow the same path, although no announcements have been made.

59. In a much-cited potential use-case published by Liberty, an airline acts as an 'Identity Provider (IdP)' and is part of a 'Circle of Trust (CoT)' in which car-hire

and hotel companies act as 'Service Providers (SPs)'. Instead of maintaining separate username-passwords for each entity, an individual can choose to 'federate' his identities, and enjoy single-sign-on – using the IdP username-password – to all entities within the CoT. In this example, it is clear that the CoT is under control of the IdP, and that the individual is constrained to use SSO, and related attribute transfer services, only for SPs that have some form of marketing agreement with the IdP. In technical terms, the Liberty specifications also support trust relationships across multiple CoTs, but questions remain about the business and governance models for such implementations.

60. As yet there have been few large-scale consumer-oriented deployments of the Liberty standards. Many in the identity community believe that this is because the Liberty members are generally large organisations who are unwilling, for commercial reasons, to cede control of personal data to the data subject. Their mindset appears to be primarily one of 'organisation-centric identity management', as defined earlier, and thus they have not backed any proposals to create general purpose infrastructure for user-centric IdM. And, since both Liberty and OASIS are standardisation organisations, neither can lead proposals in its own right.

61. Note, however, that Liberty now emphasises that its standards can also be used to enable user-centric implementations, and the design principles allow an IdP to offer the individual the ability to specify conditions and filters which control the release of her personal data to other entities in a CoT. The organisation is also supporting initiatives to develop interoperability with other IdM approaches, e.g. CardSpace and OpenID, and to create a governance framework.

Shibboleth

62. Shibboleth [SHIBB] emerged from the Internet2 grouping of universities in the USA as a solution to the problem of how students and academics of one university could be given easy access to web resources belonging to other organisations – such as other academic institutions and publishers of academic journals. Using Liberty terminology, a student's home university acts as the Identity Provider, and resource providers are the Service Providers. All the parties negotiate trust agreements within a Circle of Trust, which is sometimes called a federation.

63. Shibboleth lies part way between the organisation-centric and user-centric styles of IdM: its architects placed great emphasis on the need to protect individual privacy, making use of one-time pseudonyms (or 'handles') for identification across domains, rather than persistent unique identifiers. As a result, Shibboleth is the first large-scale framework that allows an individual to show attributes (such as 'I am a student of this university') to a third party anonymously, i.e. without also disclosing any permanent identifier. This is a likely to be a key principle for future systems.

64. Following pilots financed by JISC[1] and BECTA[2], Shibboleth is now being implemented in the UK education sector, as well as in a number of other

[1] JISC stands for the Joint Information Systems Committee, and is the organisation that funds common IT development and infrastructural needs for the UK Higher Education sector. See http://www.jisc.ac.uk

[2] BECTA is a public body responsible for ICT strategy and implementation in the secondary schools sector. See http://www.becta.org.uk

countries around the world. However, and although well designed for its limited purpose, Shibboleth would seem to lack the characteristics that would enable it to grow into a generalised infrastructure for user*-*centric IdM. Specifically: (i) the individual can only transfer attributes recorded by the organisation which hosts his Shibboleth account, rather than by any third party; (ii) the individual cannot choose which organisation hosts his Shibboleth account and (iii) the scheme lacks a business model that would allow different service providers to share the costs of secure authentication and permissioned attribute transfer.

… and the others

65. Apart from CardSpace, OpenID and Shibboleth, a few other initiatives deserve a mention. An updated listing can be found by searching Wikipedia for the term 'Identity 2.0' [WIK].

66. *PRIME* (Privacy and Identity Management for Europe) [PRIME] is a collaborative research project, part funded by the European Commission, and led by IBM Research, HP Labs and others. It kicked-off in 2004, and is due to finish in 2008. As far as can be told at present, the project has confirmed the feasibility of certain aspects of IdM infrastructures that are based on fundamental privacy-enhancing technologies, and has trialled their use for collaborative e-learning and in a location-based services application. The project also conducted research into the social, legal, economic and human-factors aspects of the resulting designs. As a research project, PRIME is not positioned to lead a roll-out of the technology that it has developed. However, T-Mobile plans to base its commercial location-based services system design on work done in PRIME.

67. Named for a long-tailed Tasmanian mouse, *Higgins* [HIG] is an open-source, user-centric, server-side IdM project led by a US start-up, Parity Communications inc. IBM and Novell are both contributing code. The authors understand that Parity intends to launch a brokerage service in late 2007, with one or more e-commerce sites as their customers. They hope to be paid a small fee per user per year by each site. As yet, there is no evidence of work to create a managed market of brokers, allowing an individual to pick which organisation would act as his broker for a relationship with any given service provider.

68. Dr Stefan Brands is, as a journalist neatly put it, the only person on the planet to have worked for both previous attempts to commercialize privacy: Digicash and Zero Knowledge Systems. He is an adjunct professor at McGill University, and also the president and founder of a Canadian software firm. His firm, *Credentica* [CRED] is developing software to commercialise a number of his patents, which show how PKI infrastructures can be made privacy-enhancing by the use of what are called 'blind signatures'. Credentica is best seen as a component supplier for user-centric IdM.

69. Edentity, a UK start-up, has developed an approach to user-centric IdM called *Personal Information Brokerage* [PIB] and is working to refine the technical, business and organisational models, and map out a route to market and critical mass. The company:

    - envisages that, in the long run, mobile network operators (and possibly banks) will be the organisations that provide a brokerage service, since they have right kind of relationship with the individual, and have (or should have) the capacity for secure authentication as a result of their existing activities.

- is seeking to develop the consensus required to run a pilot of PIB, probably leading with applications in the education sector.
- in terms of mindset is close to Parity Communications / Project Higgins in the USA, but has adopted a very different route to market, suited to the different situation in the UK.

Conclusion

70. Identity management is a fast-developing field, in which two distinct communities have yet to find a full accommodation. On one hand, there are teams within organisations, and within government, who have the resources to push forward with projects at a rapid pace. They tend to opt for organisation-centric schemes because the technology is mature, and the business, organisational and procurement models are broadly consistent with accepted practice within their spheres. Since organisation-centric schemes require that organisations group together into ever larger clusters for the purpose of back-office data sharing, they are generally harmful of privacy.

71. At the other extreme, there are many initiatives which – under the general heading of user-centric identity management – attempt to empower the individual to authenticate securely to multiple distinct counterparties, and then either give transaction-based permissions for the transfer of personal information to and between such counterparties. Since the individual can choose whether or not such information should include or exclude legal, or any other persistent, identifiers, he has the option to transact anonymously or pseudonymously, and so protect his privacy.

72. Because these user-centric initiatives require new organisational and business models, and do not fit well with the procurement procedures used by large organisations, they make slow progress. But they have significant potential to improve the general level of privacy by enabling the individual to become the gate-keeper for the use of his own personal information. They are also likely to assist organisations by:

    - reducing the costs of secure authentication,
    - reducing the frequency of catastrophic data protection beaches, and the associated reputational and financial risks;
    - contributing to efforts to ensure that customer records are kept up-to-date; and
    - enabling citizens also to control the flow of information sent *to* them, whether for marketing or other purposes.

73. Because of the need for consensus, on many technical and non-technical matters, among a large group of organisations providing services to individuals, it seems unlikely that infrastructure for user-centric IdM can be developed further without the active involvement of the public sector. Such involvement would act both as a focus and channel for attaining consensus and be seen as a welcome endorsement that reduces commercial investment risk. Thus it is for government to decide whether the potential benefits, as described above, justify the effort of becoming involved.

References

BOO See http://www.nhsconfidentiality.org/, which is the website of a group campaigning about the need for better privacy in the NHS. They also use the name 'The Big Opt-Out'.

CAM See http://www.identityblog.com for a listing of, and discussion about, Cameron's 7 laws of identity.

CP See http://www.everychildmatters.gov.uk/contactpoint/

CRED See http://www.credentica.com/

CS See http://cardspace.netfx3.com/ for information about MS Cardspace

HARDT See http://uk.youtube.com/watch?v=RrpajcAgR1E

HIG See http://www.eclipse.org/higgins/

ICO See 'A Surveillance Society', a report commissioned by the Information Commissioner for the 28th International Data Protection Commissioner's Conference, held in London in 2006. Available at http://www.ico.gov.uk

IPS 'Strategic Action Plan for the National Identity Scheme', published by the Home Office in December 2006. Available at http://www.identitycards.gov.uk.

ITW See http://www.itweek.co.uk/itweek/comment/2184695/openid-open-abuse

LA See http://www.projectliberty.org/

MIAP See http://www.miap.gov.uk/

OID See http://openid.net/

OII 'The Oxford Internet Survey 2007', by the Oxford Internet Institute. Available at http://www.oii.ox.ac.uk/microsites/oxis/events/launch2007.cfm

PIB See, for example, the outline description of PIB provided by Edentity in a submission to the recent House of Lords S&T Committee enquiry into Personal Internet Security. Available from Edentity Ltd. http://www.edentity.co.uk

PRIME See https://www.prime-project.eu/

SAML See http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=security

SHIBB See http://shibboleth.internet2.edu/

TRGD See http://www.trustguide.org.uk

WIK See http://en.wikipedia.org/wiki/Identity_2.0
