consult hyperion
securing tomorrow's transactions

# Digital Identity: Issue Analysis


CHyp doc ref: PRJ.1578
Version: 1.6
Date: 8th June 2016
Authors: Carly Nyst
Steve Pannifer
Edgar Whitley
Paul Makin
Approved: Dave Birch

*While Omidyar Network is pleased to sponsor this report, the conclusions, opinions, or points of view expressed in the report are those of the authors and do not necessarily represent the views of Omidyar Network.*

# REVISION HISTORY

| Version | Date | Author | Detail |
|---|---|---|---|
| 1.6 | 15th June 2016 | Paul Makin | Minor updates following further review. |
| 1.5 | 23rd May 2016 | Paul Makin | Final updates before publication. |
| 1.4 | 13th April 2016 | Paul Makin | Updates to reflect new information about Aadhaar. |
| 1.3 | 5th April 2016 | Steve Pannifer | Minor typo correction. |
| 1.2 | 10th March 2016 | Paul Makin, Steve Pannifer, Carly Nyst, Edgar Whitley | Final version – incorporating comments from ON and with Executive Summary now extracted into a separate document. |
| 1.1 | 22nd January 2016 | Paul Makin, Steve Pannifer, Carly Nyst, Edgar Whitley | Updated after further ON review, and comments received from external reviewers. |
| 1.0 | 22nd January 2016 | Paul Makin, Steve Pannifer, Carly Nyst, Edgar Whitley | First complete version. |
| 0.95 | 22nd December 2015 | Paul Makin, Steve Pannifer, Carly Nyst, Edgar Whitley | Second interim version shared with ON. |
| 0.9 | 1st December 2015 | Paul Makin, Steve Pannifer, Carly Nyst, Edgar Whitley | First interim version shared with ON. |

# CONTENTS

# 1 INTRODUCTION

## 1.1 The link between digital identity and privacy

Digital identity is widely recognised as being of strategic importance to the future of digital services. It has the potential to enable inclusive access to services that can be delivered in personalised and convenient ways.

There are several classes of digital identity system including:

- **Foundational**: a core digital identity, created out of a national identity scheme or similar, which is based on the formal establishment of identity and enables a wide variety of services.
- **Functional**: a digital identity normally derived from the foundational identity, which is used to address the specific needs of an individual sector, such as healthcare or banking.
- **Transactional**: a digital identity again derived from the foundational identity, which is intended to ease the conduct of transactions (either face to face or across the Internet), across multiple sectors.

Digital identity is concerned with personal data and providing mechanisms for that data to be asserted and verified in the context of digital services and transactions. A critical aspect of any digital identity scheme, especially a foundational system that is intended to be broad in application, is privacy – how to protect data to ensure the privacy of the individual.

This report considers the privacy aspects of digital identity schemes, especially those that are foundational and intended to be large scale (national or international). It considers a range of digital identity models already being pursued by governments and the private sector. The privacy characteristics of the models are explored through examining the privacy impacts and risks of particular approaches and looking at potential mitigation strategies.

There are often other important concerns when building a digital identity system such as user experience, technical interoperability and a viable commercial business model. Such things can be crucial if a digital identity system is to be successful. These parallel concerns can however conflict with privacy. The report considers these, especially focusing on where trade-offs with privacy may be necessary to achieve a successful approach.

## 1.2 Document structure

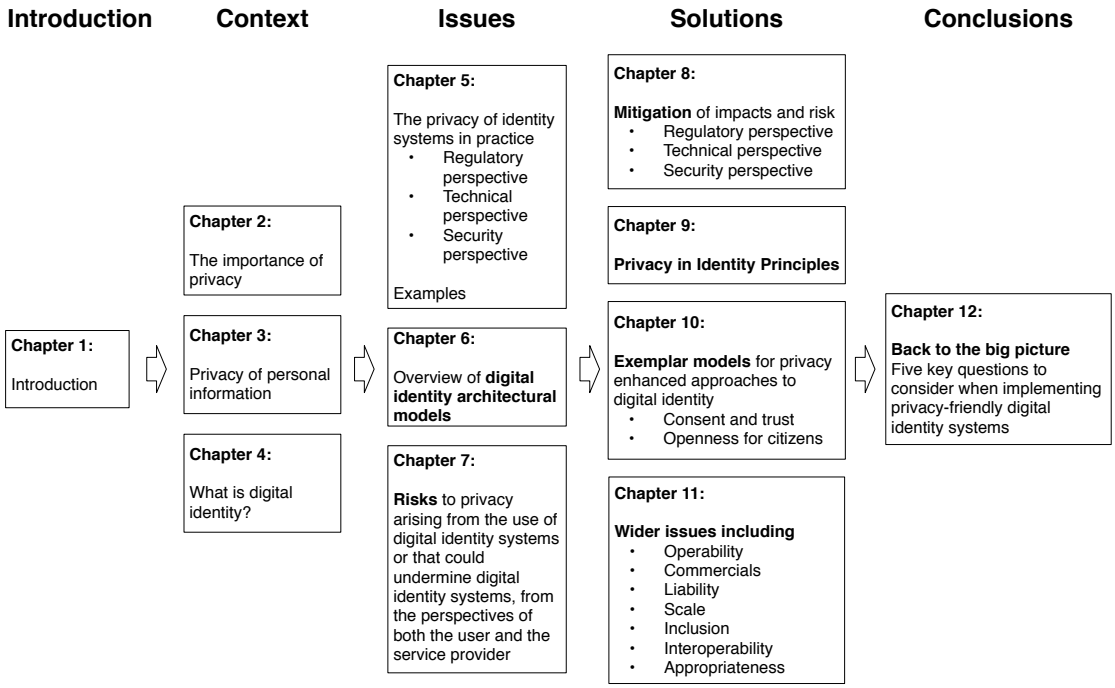



*Figure 1: Document structure*

Figure 1 shows the structure of this document.

- Chapters 2 to 4 provide background material including describing the legal basis for privacy as it pertains to personal data, describing what a digital identity scheme is and describing the primary models that exist today.
- Chapters 5 to 7 consider the privacy of digital identity schemes in practice, describe the range of digital identity architectural models we have identified, and assess the potential privacy risks in these models.
- Chapters 8 to 11 look at how the privacy risks arising from digital identity schemes can be mitigated and build on the evidence set out in the document to define a set of Digital Identity Privacy Principles (DIPPs) against which digital identity schemes can be evaluated. These are then supplemented with descriptions of good models, together with consideration of the wider issues including potential trade-offs that may need to be made with privacy.
- Chapter 12 concludes the document by reflecting on the key questions to consider when seeking to implement a privacy friendly digital identity system.

## 1.3 Acknowledgments

The authors would like to acknowledge the assistance of the following people in the preparation of this report:

- Professor Graham Greenleaf, who provided valuable input as an external reviewer;
- The entire team at Omidyar Network, and in particular Eshanthi Ranasinghe.

# 2 PRIVACY

## 2.1 The importance of privacy

Privacy is often presented as one half of a dichotomy: in conflict with security, an impediment to technological advancement, or an alternative to convenience and effectiveness. Viewed through the lens of particular political and cultural debates, privacy is a barrier: even reference to it will tend to restrain public and corporate actors from acting for the greater good and ensuring the effectiveness of the free market.

Yet the essence of privacy is to enable, rather than restrain. Privacy enables individuals to develop autonomously, independent of interference, and to fully realise their human dignity. Approaching privacy from the perspective of false dichotomies not only undervalues the important functional role it plays, it ensures this critical societal value remains locked in constant conflict with other societal imperatives. Rather, if we see privacy for what it is – a fundamental right that empowers individuals and gives them control over decisions made about them – we can see the functional role it plays in any democratic society.

Privacy is internationally recognized as a fundamental human right. It has its foundations in the constitutions of more than 100 countries, in numerous regional and international treaties;[1] and in the jurisprudence of courts across the democratic world. Privacy is at the heart of the most basic understandings of human dignity – the ability to make autonomous choices about our lives and relationships, without outside interference or intimidation, is central to who we are as human beings. Autonomy is not just about the subjective capacity of an individual to make a decision, but also about having the external social, political and technological conditions that make such a decision possible.[2] Privacy confers those external conditions. As private autonomy is a key component of public life and debate, privacy is not only a social value, but also a public good.[3] It also acts as a critical shield for individuals, protecting them from government and corporate intrusion into their homes, communications, opinions, beliefs, identities and bodies.[4]

Thus, the right to privacy is responsible for ensuring women in the United States have access to abortion;[5] innocent individuals' DNA is not routinely kept on police databases in the United Kingdom,[6] and employees have a number of safeguards against unfair dismissal in Europe.[7] It has also had a considerable influence on modern identity systems, particularly in Europe. The experience of identity cards in Nazi Germany, for example, engendered in German society a strong attachment to privacy rights. As technology has advanced and biometrics have become integrated into identity systems, particular concerns about the relationship between biometric identity systems and privacy have arisen, as evidenced by the Council of Europe's 2011

[1] See, for example, the Universal Declaration on Human Rights, Art. 12; the International Covenant on Civil and Political Rights, Art. 17; the European Convention for the Protection of Human Rights and Fundamental Freedoms, Art. 8; and the American Convention on Human Rights, Art. 11.
[2] Beate Roessler, The Value of Privacy (Cambridge: Polity Press, 2005), 62.
[3] Jurgen Habermas, Structural Transformation of the Public Sphere, (Cambridge: Polity Press, 1994).
[4] Alan Westin, 'Privacy and Freedom', Scribner, June 1967.
[5] Pursuant to the landmark US Supreme Court decision in Roe v Wade 410 U.S. 113(1973)
[6] S and Marper v United Kingdom [2008] ECHR 1581
[7] Özpınar v. Turkey - 20999/04. Judgment 19.10.2010

resolution on *The need for a global consideration of the human rights implications of biometrics*.[8]

International human rights law pertaining to the right to privacy grew out of the experiences of fascism in Europe; the Universal Declaration of Human Rights (UDHR) was drafted from 1946, as Europe was in tatters and the world was reeling from the horrors of the Holocaust. The UDHR was an attempt to draw a line in the sand, to establish the fundamental principles that define what it is to be human, against which all future governments would be measured. There was never any question that privacy, and its role as placing a critical check on state power, would be recognised.[9]

Article 12 of the UDHR enshrines the protection from unlawful or arbitrary interferences with privacy, family, home, and correspondence, as well as protection from attacks on honour and reputation. As a declaration, distinct from a treaty, the UDHR does not have binding legal value in and of itself; however, it has acquired the status of what in international law is termed "binding customary international law", meaning that it is considered as persuasive in determinations of the International Court of Justice and the United Nations. The UDHR was later complemented by two binding treaties, the International Covenant on Civil and Political Rights (ICCPR) and the International Covenant on Economic, Social and Cultural Rights (ICESCR), the first of which enshrines the right to privacy (in Article 17) in identical language to the UDHR, and to which 168 States are party. The right to privacy also became enshrined in the 1989 Convention on the Rights of the Child (Article 16), and the 1990 International Convention on the Protection of All Migrant Workers and Members of Their Families (Article 14). At the regional level, the right to privacy is protected by the 1950 European Convention on Human Rights (Article 8), the 2000 European Union Charter of Fundamental Rights (Article 7) and the 1969 American Convention on Human Rights (Article 11).

The right to privacy is enshrined in various forms in the constitutions of more than 100 countries worldwide; some research suggests up to 169 constitutions contain provisions related to privacy in its various forms (for example, privacy of correspondence, the home, family, honour and reputation, etc.).[10] Many developing countries contain rights to privacy in their constitutions, from the Gambia to Liberia, Myanmar to Nauru, primarily as a result of having adopted the text of international conventions, particularly the International Covenant on Civil and Political Rights.

[8] Resolution 1797 (2011), available at http://assembly.coe.int/nw/xml/XRef/Xref-XML2HTML-en.asp?fileid=17968&lang=en

[9] William A Schabas, The Universal Declaration of Human Rights: The Travaux Preparatoires (New York: Cambridge University Press, 2013). Initially, the right was worded as a "freedom from wrongful interference." Panama presented one of the first drafts to the UN, and articulated the right as the "Freedom from unreasonable interference with his person, home, reputation, privacy, activities, and property is the right of everyone." Panama noted that 49 countries contained constitutional provisions along those lines at the time. By June 1947 the provision, influenced by the delegation of the United States, had become: "No one shall be subjected to arbitrary searches or seizures, or to unreasonable interference with his person, home, family relations, reputation, privacy, activities, or personal property. The secrecy of correspondence shall be respected." The inclusion of the reference to the secrecy of correspondence is particularly striking here, making clear that communications - at that time primarily in the form of postal mail and telegrams - were considered to be an essential element of a person's private life. It was an important recognition that it is in corresponding, by communicating, that we develop and reveal our most intimate ideas, thoughts and beliefs; that we build and maintain our relationships; that we interact with the society in which we live. Even in 1947, it was recognised that to allow interference with correspondence will not only amount to intrusion into the private sphere, it will also have far ranging implications on our ability to express ourselves, to keep confidences, to seek and impart advice. Ultimately, the search and seizure language was removed from the right to privacy in the UDHR and the provision in the final draft of the declaration, in 1948, became: "No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation." A second sentence was added at the insistence, perhaps surprisingly, of the USSR: "Everyone has the right to the protection of the law against such interference."

[10] See the Constitute project: https://www.constituteproject.org/search?lang=en&key=privacy.

In addition, in some countries, particularly in Latin America, international treaties have direct effect in domestic law, meaning that the provisions of the ICCPR are on equal footing to domestic laws and regulations.

## 2.2 Privacy of information

The right to privacy has evolved to encapsulate a right to informational privacy, or data protection. An element of privacy with increasing centrality in modern policy and legal processes, informational privacy means that individuals can control who has data about them and what decisions are made on the basis of that data.

Informational privacy rights are often given regulatory protection in the form of "data protection" laws. Data protection laws generally regulate the conditions under which public and private entities can collect, process, retain and delete personal information, making them an important mechanism for protecting informational privacy rights. In the last three decades, more than 100 countries have adopted data protection laws, an important endorsement of informational privacy rights.

However, it does not follow that, because the right to privacy is central to the maintenance of a vibrant and liberal social and political order, it is a static concept. Its content and confines remain contested,[11] subject to never-ending games of tug-of-war between individuals, governments and corporations. Innovation and change – not just in technologies, but in migration and border flows, security and conflicts, attitudes and priorities – inform and challenge our conceptions of the private and the public.[12] The continual development of new means to undermine or protect privacy gives rise to new discussions about how to contextualise it, and new questions about its salience in changing contexts.

Equally, it does not follow that the existence of constitutional protections for the right to privacy or of data protection laws in developing countries are indications that privacy is engendered and recognised as a societal good. In a number of cultures, the autonomy and functioning of the community is prioritised over the autonomy of the individual, leading some to argue that privacy may be a Western concept. On the other hand, however, a number of non-Western cultures have long traditions of protecting privacy in the context of the home and the family; the Quran, for example, contains a number of passages on the importance of privacy. South Korea is home to one of the most active Constitutional Courts in the world when it comes to privacy; the Court has recently invalidated laws which criminalise adultery and which mandate the use of real names online, on privacy grounds.[13]

Recognition of the centrality of privacy to innovation and autonomy and its role in limiting state and corporate power is increasing. A retrospective view of changing attitudes towards privacy over only the past five years illustrates vividly the increasing salience of privacy as a societal good, at a time when it is increasingly endangered. In that space of time, Google has gone from a company whose chief executive was of the opinion that "if you have something that you don't

[11] David Lyon (ed), Surveillance as Social Sorting: Privacy, Risk and Automated Discrimination (London: Routledge, 2003) 19.
[12] For example, see the Council of Europe's expert report on profiling and data mining, Application of Convention 108 to the profiling mechanism, by Jean-Marc Dinant, Christophe Lazaro, Yves Poullet, Nathalie Lefever , Antoinette Rouvroy, January 2008.
[13] See http://www.abc.net.au/news/2015-02-26/south-korea's-constitutional-court-strikes-down-adultery-law/6267024 and http://www.bbc.co.uk/news/technology-19357160 respectively.

want anyone to know, maybe you shouldn't be doing it in the first place",[14] to offering encrypted products which ensure its users' privacy is impervious from infringement by state actors.[15] Privacy has gone from being cast as an inconvenient obstacle to the development of innovative technological products, to being a key selling feature of those same products: Companies such as Apple now trumpet, rather than seek to obscure, their efforts regarding privacy.[16]

## 2.3 What is driving privacy?

Triggering this immense shift in attitudes towards privacy has been a perfect storm of events and trends: growing public experience of data-driven targeted advertising, cyberbullying and trolling on social networks; numerous large data breaches exposing the insecurity of government data management systems;[17] and the leaks by NSA whistleblower Edward Snowden revealing the global communications surveillance efforts of the US and UK and their allies. The Snowden documents detailed the extent to which governments worldwide are intercepting communications, confirmed that even political leaders are not immune from modern surveillance; and revealed that the NSA and its allies view the present legal and technological conditions as conducive to the "golden age" of surveillance.[18]

The impact of the Snowden revelations has been to force a recognition that legal restrictions had not kept pace with technological advancements. [19] Since the documents were published, ten countries[20] countries have begun or completed legislative processes to update and modernise their laws concerning the interception and surveillance of digital communications. In many cases, however, these modernisation efforts have resulted in increased, rather than reduced, state powers to conduct surveillance, and have faced meagre public opposition. Nevertheless, studies show that individuals worldwide care more about their privacy than they did before the Snowden documents were published, and almost two thirds of people are concerned about government monitoring of their online activities.[21]

[14] Eric Schmidt speaking in 2010; "Google CEO On Privacy (VIDEO): 'If You Have Something You Don't Want Anyone To Know, Maybe You Shouldn't Be Doing It'", Huffington Post, 18 arch 20110, available at http://www.huffingtonpost.com/2009/12/07/google-ceo-on-privacy-if_n_383105.html

[15] Lucian Constantin, "Google makes secure boot, full-disk encryption mandatory for some Android 6.0 devices," PCWorld, 20 October 2015, available at http://www.pcworld.com/article/2995438/android/google-makes-full-disk-encryption-and-secure-boot-mandatory-for-some-android-60-devices.html.

[16] "Apple CEO Tim Cook: 'Privacy Is A Fundamental Human Right,' NPR, 1 October 2015 available at http://www.npr.org/sections/alltechconsidered/2015/10/01/445026470/apple-ceo-tim-cook-privacy-is-a-fundamental-human-right

[17] Marina Koren, "About Those Fingerprints Stolen in the OPM Hack," The Atlantic, 23 September 2015, available at http://www.theatlantic.com/technology/archive/2015/09/opm-hack-fingerprints/406900/

[18] James Risen and Laura Poitras, "N.S.A. Report Outlined Goals for More Power," The New York Times, 22 November 2013, available at http://www.nytimes.com/2013/11/23/us/politics/nsa-report-outlined-goals-for-more-power.html

[19] In short, the Snowden documents and related litigation have revealed that the US's NSA and UK's GCHQ were relying on outdated laws which authorised the interception of "foreign" or "external" communications, which laws were drafted during a time at which communications entering the country via particular cables could reliably said to be foreign. With the advancements of the internet and by virtue of the nature of internet infrastructure, almost all communications, even wholly domestic ones, will now transit the world and be indistinguishable from "foreign" communications. It is by manipulating these definitions that the intelligence agencies have justified intercepting *all* communications.

[20] The United States, the United Kingdom, Kenya, Switzerland, France, Germany, Finland, Denmark, Netherlands and Australia.

[21] CIGI-Ipsos Global Survey on Internet Security and Trust, undertaken by the Centre for International Governance Innovation (CIGI) and conducted by global research company Ipsos, reached 23,376 Internet users in 24 countries, and was carried out between October 7, 2014 and November 12, 2014. The full results are available at https://www.cigionline.org/internet-survey.

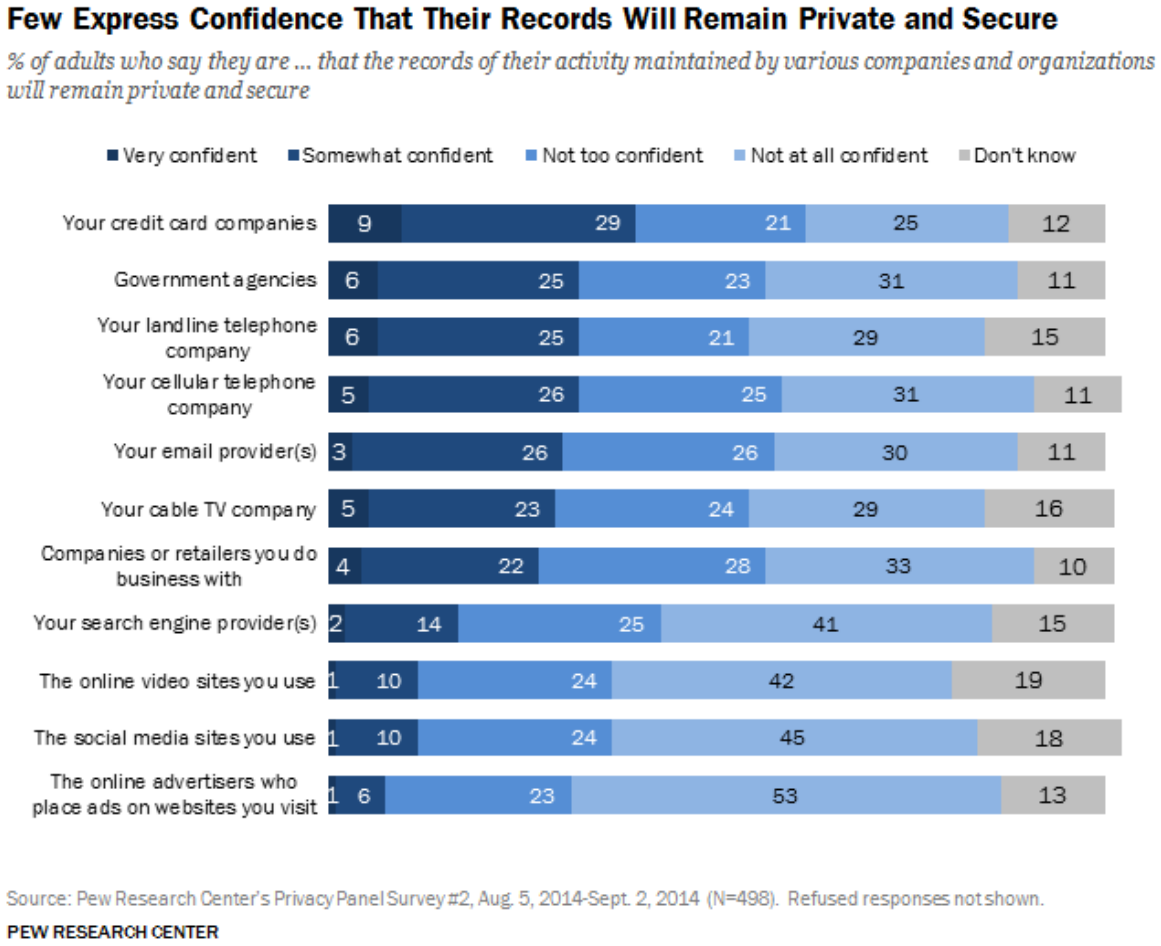



*Figure 2: Citizen Views on Privacy*

The Snowden revelations have particularly sensitised individuals to the extent of corporate data retention; in fact, studies repeatedly show that people continue to be more concerned about companies monitoring their online activities and selling that data, than they are about state surveillance.[22] This is linked to a growing distrust in the ability of corporate actors to keep data secure. Against a backdrop of generalised fear of the threats posed by cyber crime, hacking and identity theft, major data breaches such as the August 2014 hack of more than 27 million South Koreans' personal information,[23] and the (comparatively minor) TalkTalk hack in the UK in November 2015 which affected more than 150,000 of the company's customers[24] are extraordinarily costly (both financially and in terms of reputation) experiences for corporate actors, and demonstrate the importance of security risk management. The spectre of legal action looms large in the aftermath of such attacks, not to mention the reputational impact. Consequently security is now a primary driver of corporate models and behaviours, as well as a critical part of all strategic planning and risk management initiatives. The growing popularity of end-to-end encrypted services and the choice by Google and Apple to roll out full disk encryption on such devices only confirms this new reality. In addition, in the decision of the

[22] CIGI-Ipsos Global Survey on Internet Security and Trust, available at https://www.cigionline.org/internet-survey
[23] Kate Vinton, "Data Breach Bulletin: Sixteen Arrested After Allegedly Hacking Half of South Korea," Forbes, 26 August 2015, available at http://www.forbes.com/sites/katevinton/2014/08/26/data-breach-bulletin-sixteen-arrested-after-allegedly-hacking-half-of-south-korea/.
[24] "TalkTalk hack 'affected 157,000 customers'", BBC News, 6 November 2015, available at http://www.bbc.com/news/business-34743185

Court of Justice of the European Union to invalidate the Safe Harbour Agreement, in *Schrems v Data Protection Commissioner of Ireland*,[25] the Court held that companies transferring data from the EU to the US are obliged to ensure such data is guaranteed sufficient protections from intrusion. This has highlighted that the persistence of extensive state surveillance practices not only remains a threat to the privacy of a company's users, but will also shape their corporate practices.

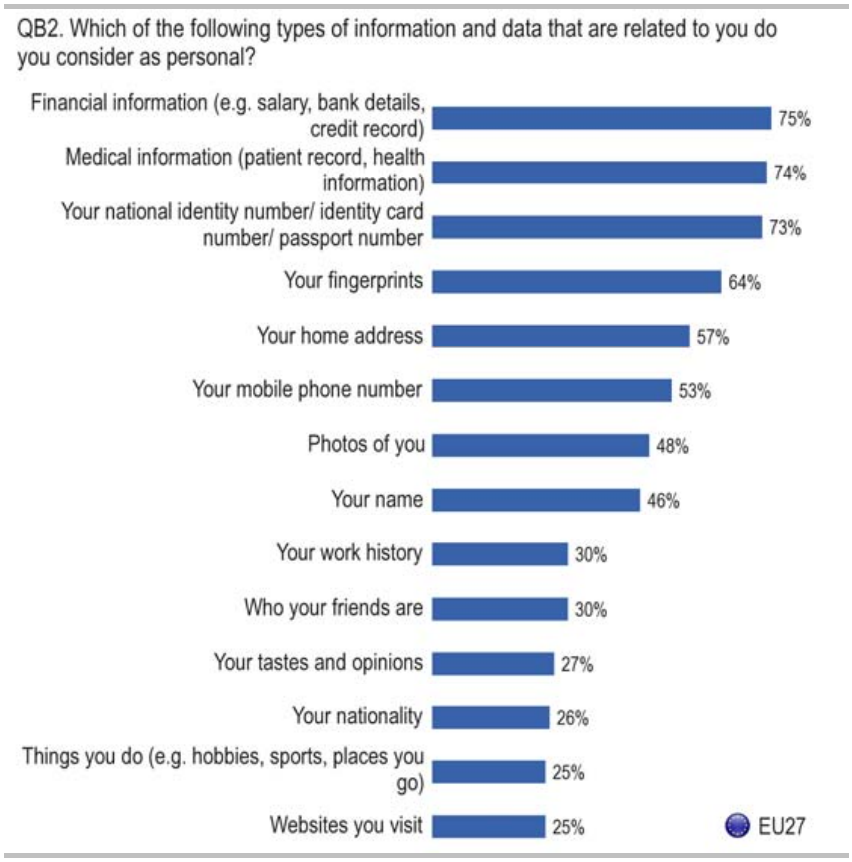



*Figure 3: European Citizen Views on Personal Data*

*(Special Eurobarometer 359 Attitudes on Data Protection and Electronic Identity in the European Union, 2011, http://ec.europa.eu/public_opinion/archives/ebs/ebs_359_en.pdf)*

## 2.4 A global perspective on privacy

In short, then, privacy is not only a fundamental right, it is a societal value of increasing salience to the public at large and the corporate sector which services them, even as attempts to debase it proliferate. This is true as much in the developing world as in the developed world. With governments struggling to manage development, security, growth and modernisation in the absence of adequate legal systems, physical infrastructure and resources, the potential for privacy to be sacrificed in favour of state control mechanisms and light corporate regulation in emerging economies cannot be overestimated. The conceptual and practical obstacles to

[25] Case C 362/14, judgement of 6 October 2015

ensuring laws and regulations keep up with rapid technology changes and expanding capabilities, obstacles that are incredibly difficult to overcome in even the most developed countries, are particularly challenging for emerging economies and democracies. Corruption, corporate influence, and weak separation of powers all dilute the strength of constitutional and legal protections. In addition, the immense challenge of addressing terrorism and other domestic and regional threats in unstable political climates manifests in watered-down safeguards for individual privacy.

Nevertheless, it is clear that privacy is increasing in recognition, importance and relevance in developing and emerging economies. In 2014 Brazil adopted a landmark piece of legislation entitled the Marco Civil da Internet, which extends considerable privacy protections to the country's internet users. 2014 was also the year that the African Union adopted the AU Convention on Cybercrime and Personal Data Protection. In China, growing awareness of the harm caused by data breaches caused the country – which still has no data protection law – to amend clause 253 of the Criminal Law in September 2015 to greatly expand the scope of privacy protection and the penalties for breaches. There is no doubt, therefore, that privacy is an issue as high on the agenda of emerging economies as it is in Europe and North America.

# 3 PRIVACY OF PERSONAL INFORMATION

Despite its aforementioned role as 21st century buzz-word and number one enemy of security, convenience and effectiveness, a widely agreed definition of privacy remains elusive. The 20th century saw numerous prominent attempts to define the right,[26] as well as a number of key initiatives to enshrine its protection in international law and domestic regulation. Since 1948, the right to privacy has been enshrined in international human rights law; however, with the technological advancements that have occurred in recent decades, privacy has come to take on new meanings and applications.

The invention and public adoption of computers forced an expansion in understanding of what privacy rights are and how they can be infringed. Rather than simply the right to be let alone, privacy came to be considered as connected with, and essential to the protection of, information. In a groundbreaking articulation of privacy rights, Colombia University Professor Alan Westin, writing in 1967, first spoke of privacy as an individual's right "to control, edit, manage, and delete information about them[selves] and decide when, how, and to what extent information is communicated to others."[27]

The 1970s saw the adoption of the first pieces of legislation relating to the protection of personal data; the first national law was adopted in Sweden in 1972, whereas the German State of Hessen adopted the world's first law on data protection in 1971; Germany adopted a federal law in 1979. The United States established the *US Secretary's Advisory Committee on Automated Personal Data Systems* which produced a 1973 report, *Records, Computers and the Rights of Citizens*, that proposed Fair Information Practice and developed a code of fair information practice for automated personal data systems. In 1974, the US adopted the Privacy Act, providing safeguards against invasion of personal privacy through the misuse of records by Federal Agencies.

[26] In common law traditions, privacy rights have their roots in two legal actions – property rights, which offered protection for one's home, papers and, eventually, image and likeness from intrusion and misappropriation, and personal rights, which originally protected against assault and nuisance to the person, and later evolved to encompass the right to reputation, giving rise to causes of action such as defamation and libel. The famous jurist Louis Brandeis, in his joint seminal 1890 paper "The right to privacy" described the development of these two legal doctrines and called for the recognition of an independent right "to be let alone", in response to "[r]ecent inventions and business methods [which] call attention to the next step which must be taken for the protection of the person […] Instantaneous photographs and newspaper enterprise have invaded the sacred precincts of private and domestic life, and numerous mechanical devices threaten to make good the prediction that "what is whispered in the closet shall be proclaimed from the housetops."( Warren and Brandeis, "The right to privacy", 4 Harvard Law Review 193-220 (1890)) Brandeis then was addressing intrusions of privacy occasioned by the press; later, in his famous dissent in the 1928 case of Olmstead v United States, which found that the US government did not require a warrant to wiretap a telephone, Brandeis addressed intrusions of privacy occasioned by the government: "Subtler and more far-reaching means of invading privacy have become available to the Government. Discovery and invention have made it possible for the Government, by means far more effective than stretching upon the rack, to obtain disclosure in court of what is whispered in the closet […] The progress of science in furnishing the Government with means of espionage is not likely to stop with wire-tapping. Ways may someday be developed by which the Government, without removing papers from secret drawers, can reproduce them in court, and by which it will be enabled to expose to a jury the most intimate occurrences of the home. Advances in the psychic and related sciences may bring means of exploring unexpressed beliefs, thoughts and emotions. "That places the liberty of every man in the hands of every petty officer" was said by James Otis of much lesser intrusions than these." (Olmstead v. United States, 277 U.S. 438 (1928) Dissent of Justice Brandeis) In formulating his own understanding of and conviction to the establishment of a self-standing right to privacy, Brandeis, having been schooled in Germany, drew from an understanding of Continental "personality rights", which had longer tradition and sought to protect both the physical and non-physical integrity of a person. The right of personality was enshrined in the German Basic Law (Grundsgericht) since 1949 and in the case law of the German Federal Constitutional Court from 1954 onwards, and guarantees as against all the world the protection of human dignity and the right to free development of the personality.

[27] Westin, Alan (1967). Privacy and Freedom. New York: Atheneum. p. 7.

Early examples of the challenges posed for privacy centred on the use of technology to develop foundational identity systems, such as in the context of national censuses. In this regard, a decision by the German Constitutional Court in 1983 was a landmark encapsulation of the challenges to privacy posed by the collection, retention and analysis of personal information, particularly in the context of technology, and on the importance of informational self-determination. The Court said that the right to privacy

> "is endangered primarily by the fact that, contrary to former practice, there is no necessity for reaching back to manually compiled cardboard-files and documents, since data concerning the personal or material relations of a specific individual can be stored without any technical restraint with the help of automatic data processing, and can be retrieved any time within seconds, regardless of the distance. Furthermore, in case of creating integrated information systems with other databases, data can be integrated into a partly or entirely complete picture of an individual, without the informed consent of the subject concerned, regarding the correctness and use of data."[28]

The right to the protection of personal information as a distinct component of the right to privacy was emerging. The United Nations Human Rights Committee, a body of human rights experts with authority to issue advisory opinions and ensure compliance of States with their treaty obligations under the International Covenant on Civil and Political Rights, followed the German Court's line of reasoning with its 1989 General Comment No. 16 on the right to privacy, distinguishing in particular the dangers posed by databases as technology advances:

> "The gathering and holding of personal information on computers, databanks and other devices, whether by public authorities or private individuals or bodies, must be regulated by law. Effective measures have to be taken by States to ensure that information concerning a person's private life does not reach the hands of persons who are not authorized by law to receive, process and use it, and is never used for purposes incompatible with the Covenant. In order to have the most effective protection of his private life, every individual should have the right to ascertain in an intelligible form, whether, and if so, what personal data is stored in automatic data files, and for what purposes. Every individual should also be able to ascertain which public [authorities] or private individuals or bodies control or may control their files. If such files contain incorrect personal data or have been collected or processed contrary to the provisions of the law, every individual should have the right to request rectification or elimination."

[28] BVerfGe 65, 1. The text is available at http://www.datenschutz-berlin.de/gesetze/sonstige/volksz.htm.

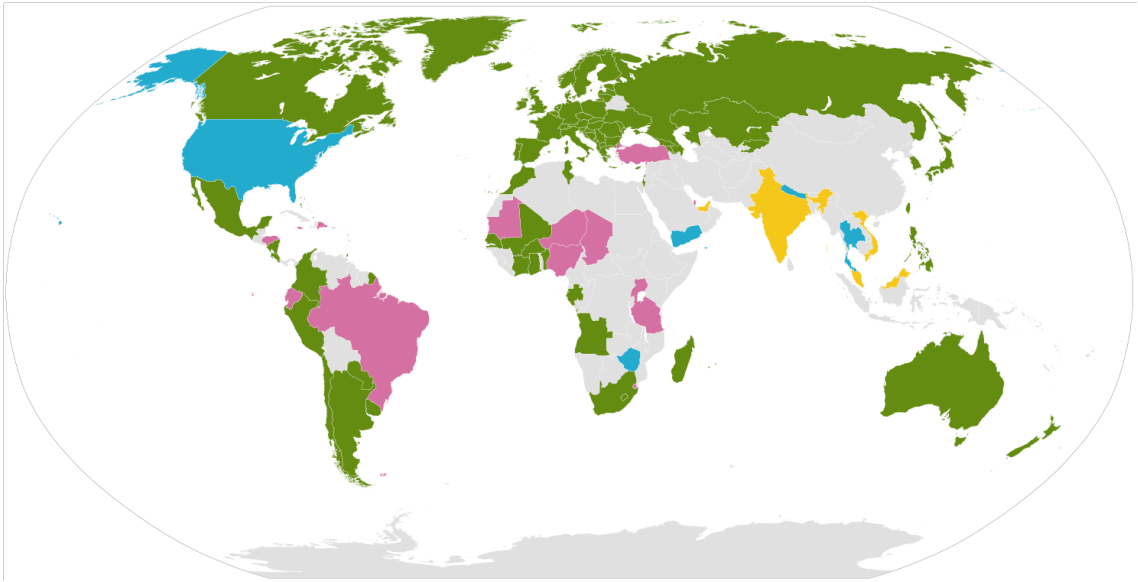



*Figure 4: Global data protection regulation*

## 3.1 The emergence of data protection principles

These sentiments were underpinned by and reinforced in the 1980 Organisation for Economic Cooperation and Development (OECD) Guidelines on the Protection of Privacy and Transborder Flows of Personal Data, the first international statement on the specific conditions under which personal information should be handled in order to ensure an individual's right to privacy is respected. The OECD's Guidelines stipulated the principles which form the basis of modern data protection law (see below).

The same principles were reflected, for the large part, in the first internationally binding instrument on the protection of personal information, the Council of Europe's Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (known as Convention 108).[29] The Convention has been signed by all 48 Council of Europe members and ratified by all but Turkey; in addition, Uruguay ratified the Convention in 2013. Four African countries are also now at the stage of acceding to Convention 108. Convention 108 adopts a broad definition of "personal data" as "any information relating to an identified or identifiable individual"[30] and "automatic processing" as the automation in whole or in part of "storage of data, carrying out of logical and/or arithmetical operations on those data, [or] their alteration, erasure, retrieval or dissemination."[31] In 2001, Convention 108 was supplemented with an Additional Protocol regarding the Automatic Processing of Personal Data regarding supervisory

[29] The text of the Convention is available at http://www.coe.int/en/web/conventions/full-list/-/conventions/rms/0900001680078b37
[30] Article 2(a)
[31] Article 2(c)

authorities and transborder data flows" (the Additional Protocol), which brought its standards up to approximately the same level as the EU Directive.[32]

## 3.2 Regulation by the European Union

It was with the creation of the European Union and its foray into law making that the protection of personal data was finally given teeth, in the form of the EU Data Protection Directive 95/46/EC (along with the E-Privacy Directive[33], the Cookie Directive [34] and the Treaty of Lisbon). The Data Protection Directive[35] (DPD), passed in 1995, requires member states to enact and enforce legislation protecting individuals with regard to the processing of personal data, while promoting data sharing within the bounds of the European Convention of Human Rights, particularly the right to privacy (Article 8). The principles it enshrines echo those in Convention 108, and it requires member states to meet minimum requirements through their choice of measures. The directive also establishes the Working Party on the Protection of Individuals with regard to the processing of personal data, also known as the Article 29 Working Party, comprised of representatives of each of the member states' data protection authorities, as well as representatives from the European Commission. The Article 29 Working Party examines and issues opinions with respect to the application of the directive.

The operation of the DPD is bolstered by the coming into force in 2009 of the Charter of Fundamental Rights of the European Union, which in addition to enshrining the right to privacy (Article 7) in the same terms as Article 8 of the European Convention of Human Rights, includes a separate and distinct right to the protection of personal data (Article 8). In recent years, Article 8 of the Charter has played an important role in the approach of the Court of Justice of the European Union (CJEU) to issues concerning privacy and data protection. It was key to two recent seminal decisions issued by the Court, that concerning the mandatory retention of telecommunications data (*Digital Rights Ireland*) and the adequacy of United States privacy protections in the context of transfer of EU data abroad (*Schrems v Data Protection Commissioner of Ireland*), both of which are discussed further below. As a general conclusion it can be said that the coming into force of the Charter, and particularly Article 7, has caused the Court to put greater emphasis on the need for independent authorisation of the processing, retention of and access to personal information.

The DPD has its limitations, not least of which is that it is two decades old and fails to address many of the current realities of and challenges to the protection of personal data. These challenges include those posed by the advent of social networking and thus the expansion of the types of personal data generated by individuals and processed by third parties; the advent of cloud computing and big data; and divergent enforcement regimes and lack of harmonisation

[32] Available at http://conventions.coe.int/treaty/en/treaties/html/181.htm.
[33] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications), 2002 O.J. (L 201) 37 [E-Privacy Directive], available at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2002:201:0037: 0047:EN:PDF.
[34] Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws, 2009 O.J. (L 337) 11 [the Cookie Directive], available at http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:337: 0011:0036:en:PDF
[35] Available at http://eur-lex.europa.eu/LexUriServ/ LexUriServ.do?uri=OJ:L:1995:281:0031:0050:EN:PDF. Directives, as distinct from regulations, are not self-executing but rather require member states to enact legislation to achieve a particular result – in this context, the protection of personal data – without dictating the means of achieving that result.

across respective member states. [36] In order to overcome these and other issues, the European Union began, in 2012, to negotiate a Data Protection Regulation which would have direct effect in member states. The Regulation was finalised and adopted by the Parliament on 15 December 2015, although its provisions won't come into force until 2018. Additionally, the EU is negotiating a proposal for a new Directive to replace the current Council Framework Decision 2008/977/JHA for data protection in the area of criminal law enforcement.

## 3.3 Following the leader? Data protection outside Europe

Today, there are more than 100 national data privacy laws around the world, more than half of which are from outside Europe.[37] Many of these laws closely mimic European standards: the DPD has had an impressive influence on the development of data protection law in countries outside of Europe, primarily due to the restrictions preventing the transfer of data outside the EU where the third country does not have an "adequate" level of data protection. Although "adequacy" is a broad concept and can be achieved through a number of routes, the normative impact of this requirement has been to create incentives for third countries to bolster the levels of data protection they provide, in order to ensure they contribute to be a hospitable environment for industry. This is particularly the case with respect to countries which are business process outsourcing destinations; the Philippines, for example, recently adopted the Data Privacy Act 2012, heavily influenced by the European DPD (although it has not yet come into effect). It is possible to identify 33 non-European countries which have legislated for data protection frameworks that substantially incorporate the higher protections of the DPD, rather than the broader principles enshrined in the OECD Guidelines.[38]

In addition, there are a number of other regional instruments which enshrine data protection rights and principles. In Asia, influenced by the OECD Guidelines and Convention 108, the Asia-Pacific Economic Cooperation (APEC) adopted a non-binding Privacy Framework in 2004. Ten Asian countries possess national data protection legislative frameworks; neither India nor China is among them. In addition, the ASEAN Human Rights Declaration, adopted in 2012, while non-binding in nature, includes particular reference to the protection of personal data to its provision on the right to privacy.

The African Charter of Human and Peoples' Rights notably omits to include a right to privacy. However, in June 2014, the African Union adopted the African Union Convention on Cyber Security and Personal Data Protection. The Convention has received a mixed response, with many criticising the weak provisions and vague definitions contained therein. Moreover, no government has yet ratified the treaty, which will only come into force after 15 states have ratified.[39]

Although a number of countries in the Americas already have national data protection laws, and Argentina and Uruguay enjoy adequacy status vis a vis the EU DPD, most countries in the

[36] For a more fulsome exploration of these challenges, see Marc Rotenberg and David Jacobs, "Updating the Law of Information Privacy: The New Framework of the European Union", 36 Harvard Journal of Law and Public Policy 2, (2012) 605.

[37] Graham Greenleaf, Asian Data Privacy Laws (Oxford, Oxford University Press: 2014), 55. For details about each of the domestic frameworks, see BakerHostetler, 2015 International Compendium of Data Privacy Laws, available at http://www.bakerlaw.com/files/Uploads/Documents/Data%20Breach%20documents/International-Compendium-of-Data-Privacy-Laws.pdf

[38] Graham Greenleaf, Asian Data Privacy Laws (Oxford, Oxford University Press: 2014), 57.

39 Henry Roigas, "Mixed Feedback on the 'African Union Convention on Cyber Security and Personal Data Protection'", NATO Cooperative Cyber Defence Centre of Excellence, available at https://ccdcoe.org/mixed-feedback-african-union-convention-cyber-security-and-personal-data-protection.html.

region, including Brazil, have no comprehensive regulation in this area[40]. In the United States there remains no comprehensive data protection framework or federal regulatory approach; rather, sectoral and self-regulation abounds.

The Organisation of American States has only recently made concrete moves in an attempt to create coherence in data protection across the region and improve protections for privacy. The Inter-American Juridical Committee presented a report[41] to the OAS Permanent Council on 31 March 2015 which details the state of data protection debates internationally, and provides a draft model law on personal data protection. The report concludes that the OAS should approach coherence in the region by proposing legislative guidelines based on 12 principles previously espoused by the Committee in 2012 (CJI/RES.186(LXXX-0/12), rather than agree on the exact wording of a common law, following the approach of the DPD.

By no means should the proliferation of laws be taken to mean that data protection practices in developing countries are equivalent to those in Europe. A combination of lack of capacity on the part of companies and government bodies to comply with legislation, poor enforcement mechanisms, and a general absence of training and understanding about why the protection of personal information is important continues to prevent developing countries from raising the level of data protection in practice to a level even close to that in Europe. It is important to note that complying with data protection may be expensive, complex and requires not only detailed legal understanding, but rigorous processes and procedures to ensure compliance. The example of Peru is apposite (see section 5.2.10). Although the country is generally regarded as one with a comprehensive data protection law, enforcement can be problematic. The data protection law came into effect in July 2013, however by July 2014 only 90 Personal Data Filing Systems had been registered in the whole of the country, despite the law placing an obligation on every data controller (including both corporate and government entities) to register such a system[42].

[40] Brazil's recently adopted Marco Civil da Internet does contain some isolated protections, however.

41 Inter-American Juridical Committee, Privacy and Data Protection, 26 March 2015, available at http://www.oas.org/en/sla/dil/docs/CJI-doc_474-15_rev2.pdf

[42] Eliana Lesem, "Peru: Recent Updates to Peruvian Data Protection and Privacy Law," *Mondaq,* 28 July 2014, available at http://www.mondaq.com/x/330672/Data+Protection+Privacy/RECENT+UPDATES+TO+PERUVIAN+DATA+PROTECTION+PRIVACY+LAW

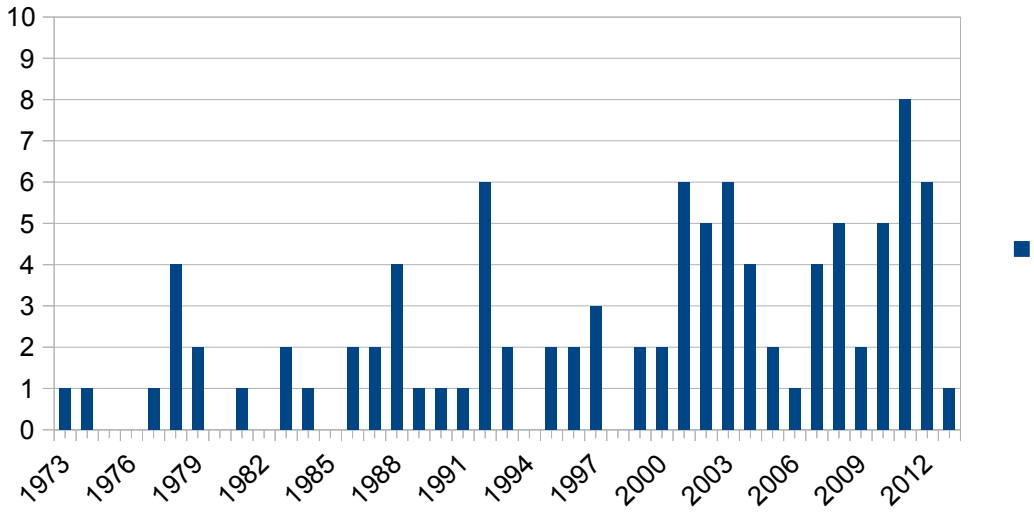



*Figure 5: Number of national data protection laws adopted annually*

## 3.4 Current challenges to informational privacy

Today, the context in which privacy and data protection are being discussed, protected and threatened is rapidly changing; even in the past five years, jurisprudence on and understandings of privacy have advanced dramatically. A number of particular interwoven issues have had particular influence as legislatures, courts, the media and the global public have engaged in a dialogue about what meaning to give "the right to privacy in the digital age".

### 3.4.1 Mass surveillance and data retention

It is trite and entirely insufficient to say that the internet has "revolutionised" how we communicate, and thus how governments conduct communications surveillance. Such a phrase does not adequately encapsulate the fundamental rupture between the nature of communications prior to the advancements of the digital era, and today's reality. Not only have communications been completely transformed in form, scope, speed and reach of communications, the definition of the act of communicating has been altered – we now communicate not only with other individuals or with our local community, but with the world at large, with our devices, with cell towers, with foreign-based internet platforms and servers, with the cloud. Simply by using a device we are transmitting – communicating – private data about ourselves, our location and our correspondence to untold entities. This private data travels across both private and public spaces; the major platforms on and services through which we communicate are operated by private companies utilising privately and publicly owned telecommunications infrastructure and complying with State licensing and spectrum allocation regulations across numerous jurisdictions. Traditional distinctions between public and private break down in this context, creating very real challenges to determining that which is fairly at the disposal of State and corporate actors, and that which must be treated as private personal information and afforded the requisite protections.

As these distinctions have disintegrated, so too has another dichotomy: that between national security, on the one hand, and law enforcement, on the other. Traditionally, the former was

outward looking – to military threats, inter-state espionage, and potential acts of aggression by State actors. Law enforcement, on the other hand, was directed internally, at the prevention, detection, investigation and prosecution of crime and disorder. Yet with the watershed events of September 2001, the replacement of the Cold War narrative with a new international discourse of the "war on terror", and the global change in security priorities, promoted by the United States and its allies, towards greater investment in the suppression of both domestic and foreign extremism, national security and law enforcement objectives have increasingly coalesced, and tactics previously in the realm of national security actors – chief among them, intelligence gathering – have increasingly become the preserve of police.

The simultaneous occurrence of these phenomena – the rapid technical advancements in communications, on the one hand, and the birth of a new global security paradigm, on the other, have motivated an approach to communications surveillance that eschews traditional understandings of targeted interception that is lawful, necessary and proportionate. Instead it has proved to be fertile breeding ground that has allowed a particular conceptualisation of modern communications surveillance to thrive: one which says that all acts of digital communications are potentially necessary pieces of an unwieldy security puzzle that can only be solved by collecting every piece. Or, to use a more popular analogy: that effective law enforcement and the protection of national security requires the identification of needles in a haystack, and the only way to so identify the needles is to collect every piece of hay available. This mindset has given birth to what are commonly known as mass surveillance programmes (although they have been rebranded by numerous key States as bulk collection or bulk interception programmes).

Mass surveillance capabilities form a key part of American and British surveillance apparatuses, as well as those of Australia, New Zealand, Canada, Germany, France, Switzerland, Sweden, the Netherlands and Denmark. Mass interception systems can also be purchased on the private market; French companies Qosmos and Amesys famously sold such technology to Gaddafi's Libya in the mid 2000s.[43] That mass surveillance measures are in place is not avowed by most countries (Britain recently introduced legislation containing powers to commit "bulk interception" with the caveat that "this is not mass surveillance"), nor is it necessarily provided for in domestic statutes, although a recent spate of legislative reform across Europe threatens to provide ostensible legal cover for such activities.

Debate still rages about the implications of mass surveillance programmes for privacy rights. States say they are simply responding to changes in technology and the global nature of national security threats; United Nations human rights experts, such UN Special Rapporteur on protecting human rights while countering terrorism, Ben Emmerson QC, disagree:

> "In the view of the Special Rapporteur, the very existence of mass surveillance programmes constitutes a potentially disproportionate interference with the right to privacy. Shortly put, it is incompatible with existing concepts of privacy for States to collect all communications or metadata all the time indiscriminately. The very essence of the right to the privacy of communication is that infringements must be exceptional, and justified on a case-by-case basis."[44]

[43] FIDH, "Amesys and Qosmos targeted by the judiciary: is there a new law on the horizon?", 18 June 2013, available at https://www.fidh.org/en/region/europe-central-asia/france/amesys-and-qosmos-targeted-by-the-judiciary-is-there-a-new-law-on-the-13966

[44] (2014, A/69/397) at [16].

In the period since the first of the Snowden documents were published, the United Nations has adopted a number of resolutions and produced a number of reports concerning mass and extraterritorial surveillance. [45] The question of whether mass surveillance is a justifiable interference with the right to privacy has been debated in numerous court cases in the United States [46] and Europe, [47] and the European Court of Human Rights is expected to rule on the issue in 2016.

In making its decision, the Court will no doubt look to recent legal advancements regarding mass/bulk metadata retention and collection (as distinct from the interception of the content of communications). In 2015, the US Court of Appeals for the Second Circuit declared the NSA's bulk telephone records collection programme unconstitutional, [48] and US Congress passed legislation, the USA FREEDOM Act, restricting the programme. Meanwhile, in April 2014 the Court of Justice of the European Union (CJEU) invalidated the European Data Retention Directive, which required member states to compel communications services providers to retain metadata records for up to two years for law enforcement and national security purposes, finding it violated Articles 7 and 8 of the Charter of Fundamental Rights. [49] These decisions reflect increasingly greater recognition of the dangers inherent in collecting and retaining data in bulk, and the need to minimise such data. Moreover, they suggest conceptualisations of the content of individual privacy rights continue to change and expand, as technology improves and the ability to extract private information out of even isolated and disparate pieces of seemingly innocuous data advances.

However, it remains difficult to square court decisions, UN resolutions and government regulation promoting the right to privacy and calling for its continued protection, on the one hand, and the increase in mass surveillance and other monitoring systems, on the other. The recent CJEU Safe Harbour decision (see below) is illustrative of the often hypocritical approaches in this respect; the European Court criticised the mass surveillance laws in the United States, concluding that their existence undermines fundamental human rights. Yet at least three European States (the UK, France, and Germany) practice mass surveillance in the same form as the US, and another four (Denmark, Switzerland, Finland and Netherlands) are in the process of updating their legal frameworks to enable them to do the same. At the same time, companies across Europe are building mass surveillance systems and selling them to a range of other states. In some respects, therefore, the increase in rhetorical commitments to privacy can be seen as a demonstration of a “do as I say, not as I do” attitude on behalf of certain western states. In another view, it could be said that when it comes to commercial matters, States are happy to regulate for the strict protection of privacy, but when it comes to security and crime prevention, the threat of terrorism is sufficient to justify even the most serious intrusions.

[45] Report of the UN Special Rapporteur on freedom of expression, A/HRC/23/40, 17 April 2013; Resolution A/C.3/68/L.45, November 2013; Report of the High Commissioner for Human Rights, The right to privacy in the digital age, A/HRC/27/37, 30 June 2014; Resolution A/C.3/69/L.26/Rev.1, 19 November 2014; Resolution A/HRC/28/L.27, 24 March 2015.
[46] For example, Wikimedia v NSA filed in May 2015: the plaintiffs' complaint is available at https://www.aclu.org/legal-document/wikimedia-v-nsa-complaint.
[47] For example, Liberty & Ors v GCHQ, [2014] UKIPTrib 13_77-H (5 December 2014)
[48] ACLU v Clapper, US Court of Appeals for the Second Circuit, 7 May 2015
[49] Digital Rights Ireland v Ireland & Ors, Court of Justice of the European Union, 8 April 2014

### 3.4.2 Cross-jurisdictional data transfers

A recent decision of the CJEU confirms that the existence of mass surveillance programmes impacts on the ability of corporate entities to transfer personal data across borders under the European legal framework. In *Schrems v Data Protection Commissioner of Ireland*, [50] (also known as the Safe Harbour decision) the CJEU addresses the requirement in the European DPD that companies and governments collecting data in Europe must ensure that any jurisdiction to which they transfer that data (including by keeping data in "the cloud" in circumstances in which physical servers are located outside of Europe) has in place adequate protections for privacy. The Court held that any third country to which data is transferred must "ensure, by reason if its domestic law or its international commitments, a level of protection of fundamental rights and freedoms that is essentially equivalent" to that which is guaranteed in the EU under the DPD and the Charter. [51] In that regard, the Court said, the existence of a mass surveillance programme would establish that the requisite levels of rights protection are not met by the third country, given that "legislation permitting the public authorities to have access on a generalised basis to the content of electronic communications must be regarded as compromising the essence of the fundamental right to respect for private life, as guaranteed by Article 7 of the Charter" [94].

The *Schrems* decision is recent and its full implications are not yet known; in essence the Court has precluded the further transfer of data to the US until legislative change is effected to preclude the type of mass surveillance programmes that are being operated under the auspices of the Foreign Intelligence Surveillance Act and were revealed by NSA whistleblower Edward Snowden. The decision comes amidst the negotiation of the new European Data Protection Regulation and a potential successor to the Safe Harbour Agreement (which the *Schrems* decision invalidated), and will have implications for the outcome of that process. But its impact will likely be felt beyond the US and EU; as cross-jurisdictional data transfers become increasingly necessary (and increasingly a matter of trade, with the current trade negotiations of the Transatlantic Trade and Investment Partnership and the Trade in Services Agreement both speaking to data transfers), the proactive standards-setting process engaged in by the CJEU is likely to have a normative effect on the development of stronger privacy protection in jurisdictions throughout Asia, Africa and Latin America.

It will also have implications for digital identity schemes which rely upon private sector entities based in Europe, where such entities process data in Europe and use US cloud services. In such cases, the entity will not be able to further transfer data outside of Europe unless it is to a country which provides "adequate" data protection regulation, and it will not be able to use US cloud services or otherwise transfer data to the US by relying on the Safe Harbour scheme. There are, however, exceptions to this, and other ways to lawfully transfer data to the US, so this calculation will need to be made on a case by case basis.

The Safe Harbour agreement was specific to the EU Data Protection Directive, and does not have equivalence in other national data protection regimes, so there are no flow-on effects to developing countries in that sense. However, it is easy to see other countries following Europe's approach and preventing the transfer of data to the US, or to another country that employs mass surveillance infrastructure. Equally, the decision may be used as a lever to argue for data localisation practices, whereby international companies are obliged to keep data collected in

[50] Court of Justice of the European Union, 6 October 2015
[51] At [73]

one jurisdiction in a physical server within that jurisdiction. Such a requirement has already been enacted by Russia[52] and was previously mooted in Brazil.[53]

### 3.4.3 Mandatory use of identity online

In order to facilitate surveillance aims, governments are with increasing frequency looking for ways to undermine the ability of individuals to be anonymous online. Whereas encryption provides security from interference with the content of a communication, it does not guarantee the anonymity of the sender or recipient of that communication, and separate measures must be taken to mask one's identity from detection. These measures may range from the use of a pseudonym online to the use of non-registered SIM cards or of anonymisation tools such as Tor.

Recent years have seen a range of state measures requiring the mandatory use or registration of identity online, including laws requiring the use of real names by bloggers and internet commentators, the registration of SIM cards and IP addresses, and the production of identification at cybercafes, as well as mandatory retention of, and State access to, metadata. While some such measures have recently been curtailed by courts – the Supreme Court of Canada referenced the importance of enabling anonymity in its decision in *R v Spencer* [54] holding that law enforcement access to subscriber information requires judicial authorisation, and the Constitutional Court of the Republic of Korea struck down anti-anonymity laws as unconstitutional [55] – but they are proliferating in many other jurisdictions. A law currently under consideration in Brazil, Bill PL215/2015, would make it compulsory for all communications service providers (including those providing internet applications as well as telecommunications access) to collect identifying information on their users, including email addresses, telephone numbers and national identity numbers. The original version of the Bill was designed, according to the government, to establish greater rigour in prosecuting crimes against honour taking place on social media.[56]

The right to remain anonymous is one element of the right to privacy, and like privacy it is not an absolute right. However, there is little consensus in national or international laws as to scope of the right to remain anonymous. Whereas there are strong protections in US law enabling individuals to communicate anonymously, [57] in Brazil, for example, the Constitution forbids anonymity. A recent report by the United Nations Special Rapporteur on freedom of expression, on anonymity and encryption, holds that any restrictions on anonymity must comply with a strict test of lawfulness, necessity and proportionality. [58] The Committee of Ministers of the Council of Europe adopted a Declaration on freedom of communication on the Internet which establishes anonymity as a central principle of freedom of communication, declaring that "in order to ensure protection against online surveillance and to enhance the free expression of information and ideas, member states should respect the will of users of the Internet not to disclose their identity."[59]

[52] http://www2.deloitte.com/be/en/pages/risk/articles/data-localisation-requirement-russia.html
[53] http://www.law360.com/articles/520198/brazil-nixes-data-localization-mandate-from-internet-bill
[54] [2014] 2 SCR 212, available at http://scc-csc.lexum.com/scc-csc/scc-csc/en/item/14233/index.do
[55] Decision 2010 Hun-Ma 47, 252 (consolidated) announced 28 August 2012.
[56] Danny O'Brien, "Brazil's Politicians Aim to Add Mandatory Real Names and a Right to Erase History to the Marco Civil", 14 October 2015, available at https://www.eff.org/deeplinks/2015/10/brazils-terrible-pl215
[57] McIntyre v. Ohio Elections Commission 514 U.S., at 342
[58] A/HRC/29/32
[59] Principle 7 - https://wcd.coe.int/ViewDoc.jsp?id=37031

There has been some recent debate about the extent to which the mandatory use of or registration of identity online can be used to support the right to privacy, while undermining enjoyment of freedom of expression. The recent decision of the Grand Chamber of the European Court of Human Rights in the case of *Delphi v Estonia*[60] has called the interconnectivity of these two rights into question, by holding that a website provider is under an obligation to be able to identify users posting comments on the site in case those users make defamatory or other unlawful comments. The decision has been met with considerable criticism, yet it reflects an arguably growing desire of both private and public actors to ensure that internet users can be identified in certain circumstances.

### 3.4.4 Cyber security

Competing for attention with terrorism in the security space are the threats posed, to critical national infrastructure, in the field of cyber security. To give an indication of the scale, the Australian Cyber Security Centre recently released a report noting that in 2014, the Australian Cyber Emergency Response Team responded to 11,073 cyber security incidents affecting Australian businesses, 153 of which involved systems of national interest, critical infrastructure and government. [61] The British Home Secretary, in presenting a new surveillance law to the British parliament in November 2015, stated that 90 per cent of large organisations in the UK suffered a security breach in the preceding year.[62]

The challenge of ensuring cyber security has led to a number of efforts on the regional and international stage, including the creation in 2015 of the Global Forum on Cyber Expertise, a permanent forum housed in The Hague and populated by 42 countries. In addition, the UN Group of Governmental Experts on Cybersecurity is a more powerful convening of states which assesses advancements in cybersecurity policy of UN member states and is currently seeking to elaborate initial-stage standards for the maintenance of security in cyberspace. Yet, save for the 2004 Budapest Convention on Cybercrime, which pertains to the domestic regulation of criminal behaviour online, there remains little international agreement as to the legal regime applicable to international cyber security threats.

From the perspective of privacy and data protection, the emerging threats in the cyber security space have three prominent implications: first, they require rigorous and innovative steps on the part of private and public sector entities to secure the private data they hold; second, they provide additional incentives for those entities to minimise the amount of data they hold in the first place, which has direct and immediate benefits for individual privacy; finally, they run contrary to existing mass surveillance measures, to the extent that addressing cyber insecurities requires promoting the deployment of ubiquitous and strong encryption tools and services, even while such tools and services frustrate mass surveillance. It may be cyber security concerns, rather than any commitment to privacy and data protection, that ultimately dissuade States from mass interception and bulk data retention priorities.

[60] App. no. 64569/09, 15 June 2015
[61] Australian Cyber Security Centre, 2015 Threat Report, available at https://www.acsc.gov.au/publications/ACSC_Threat_Report_2015.pdf
[62] "Theresa May: Internet data will be recorded under new spy laws," The Telegraph, 4 November 2015, available at http://www.telegraph.co.uk/news/uknews/terrorism-in-the-uk/11974112/New-spying-powers-to-be-unveiled-by-Theresa-May-live.html

# 4 WHAT IS DIGITAL IDENTITY?

## 4.1 Digital identity is a broad term

"Digital identity" is a term that is understood in many different ways. For example, it can refer to logon and access control, to digital signatures and non-repudiation, to establishing legal identity or to using data to gain a better understanding of the consumer.

In reality digital identity can encompass all of these things. Commercially it is usually concerned with building more trusted relationships with customers that allow:

- Better knowledge of the customer – allowing more relevant and tailored services to be delivered.
- Security with less friction – using a range of digital technologies to allow customers to use services with confidence whilst not placing unnecessary barriers in the way.

## 4.2 Inconsistent language

There is no single accepted international standard for the terminology used to describe digital identity. This section describes the main concepts and terms, and highlights the most significant ambiguities that exist in the language of identity

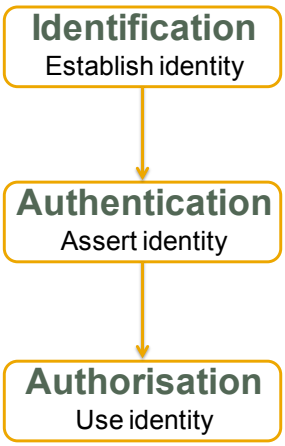



*Figure 6: What is digital identity?*

Three key concepts are core to digital identity:

- **Identification**: this is the process of *establishing* information about an individual (or organisation). It may involve examining "breeder documents" such as passports and birth certificates, consulting alternative sources of data to corroborate the identity being claimed and potentially collecting biometric data from the individual.
- **Authentication**: this is the process of *asserting* an identity previously established during identification. Typically this involves presenting or using an authentication credential (that was bound to the identity during identification) to demonstrate that the individual (or organisation) owns and is in control of the digital identity being asserted.

- **Authorisation**: this is the process of *determining* what actions may be performed or services accessed on the basis of the asserted and authenticated identity.

In this paper, "digital identity" means any system (or scheme[63]) where identification, authentication and authorisation are all performed digitally.

For example, if a social security number is entered into a web site in order to gain access to a service but there is no authentication step to confirm that the social security number belongs to the person making the request, then we do not consider this to be digital identity.

It is also important not to confuse digital identity with the channel through which the service is delivered. It is possible, for example, for a digital identity to be used in a face-to-face environment. An individual could physically attend the service provider location but assert their identity through the use of a smart card or mobile device, which are digital means of authenticating the claimed identity.

In addition to the core concepts of identification, authentication and authorisation, there are several other significant terms used in the context of digital identity. Some of these terms can be used with differing meanings. Hence the following table lists these terms and indicates the range of usage.

| Term | Range of meanings |
|---|---|
| Identity | • A individual, distinguishable from other individuals within a population<br>• The core attributes associated with an individual (name, address, date of birth) |
| Attribute | • A specific data item pertaining to an individual. |
| Partial Identity | • A subset of the attributes that make up an identity (e.g. just date of birth). This includes reduction in precision in attribute data (e.g. age bracket instead of date of birth). |
| Credential | • An authentication token (e.g. smart card) used to assert identity<br>• A verifiable attribute, e.g. a digital certificate that demonstrates an entitlement or qualification |
| Binding | • The process of linking an authentication credential to an identity in order that the authentication credential can be relied upon later on as a means of asserting the identity. |
| Subject or Principal | • The party needing to assert identity |

[63] For example, UK Verify is an identity scheme, rather than identity system; it specifies the identity verification process, the forms of digital identity supported, and how the various scheme participants interact. An analogy might be with a payment scheme such as Visa, which does not conduct payments, but as a payment scheme it defines standards and enables all of the parties to a transaction to work together.

| Term | Range of meanings |
|---|---|
| Relying Party | • The party needing to verify identity<br>• Commonly this term is used to refer to a service provider, such as a government department or retailer. We find this use of the term misleading as it suggests that identity and authentication is only important in one direction (the service provider needs to be able to authenticate the individual), however in reality authentication should be mutual (the individual has just as much need to be able to authenticate the service provider, that is the service provider demonstrates its identity before the individual shares data with it). |
| Level of Assurance | • A measure of the quality of the identity derived from the both the quality of the identification process and the strength of the authentication credential used when asserting the identity. |

*Table 1: Terms Used in Digital Identity*

## 4.3 The elements of identification

Whilst it is apparent that there is no formally agreed international definition of legal identity, it is the case that the term 'legal identity' is used in many contexts, and we take it to refer to that identity that is associated with an individual who it has been formally established has a right to reside in a country, with all of the rights and responsibilities that flow from that.

A digital identity is then a (usually) cryptographically-protected record that is associated with the legal identity, and which can therefore be used to assert, or provide assurance of, the identity of the holder in either face to face or remote (Internet or mobile) environments; so the digital identity might exist on a secure device, such as a smartcard or mobile phone, or in 'the cloud'. The corresponding assurance might take the form of a biometric check that the person presenting the digital identity is indeed the individual whose identity it asserts.

So in the ideal world there are two steps in the creation and issuance of a digital identity:

- **Civil registration**; the establishment of a 'foundational' identity, through the issuance of a recognised document. This is usually a birth certificate, reflecting an entry in a legal registry of birth and death events that occur in a country, as well as those that occur on 'foreign missions'.
- **Identity issuance**; a universal, multi-purpose system capable of supporting identity services for multiple stakeholders.

Each of these has its own legal basis, as discussed in the following subsections.

### 4.3.1 Establishing Identity

#### *4.3.1.1 Birth Certificates*

The birth certificate is in many senses the fundamental identity document. Indeed, Goal 16.9 of the UN's Sustainable Development Goals states:

> *By 2030, provide legal identity for all, including birth registration*

Across the developed world, the issuance of birth certificates and the registration of deaths are well-established, formal processes. These typically follow a similar pattern, in which attending midwives or other medical staff issue a 'birth notification' document to the mother (and typically also notify the registration authorities, together with any information about the mother that they hold). The parents are then required to present themselves to the public registration authorities within a set time period of perhaps one or two months, in order to formally register the birth, including notification of parental relationships and home address. It is this registration that is at the root of legal identity.

However, there are complications in many emerging economies, due to non-issuance of birth certificates; for example, according to a UNICEF report on Nigeria (2007):

> *"….in urban areas, approximately 50 percent of births are registered, while in rural areas, only about 21 percent are registered (UN July 2007). Low registration rates in Nigeria have been attributed to a number of factors, including lack of awareness of current legislation and of the importance of birth registration, limited number of registration centres, limited financial resources and a lack of effective registration infrastructures".*

In many cases, there are strong correlations between communities that are already at the fringes of society and those who lack proper birth registration. This leads to a wide range of issues[64], ranging from non-issuance of national identity cards through to problems with immigration into Western countries such as the United States, and of course financial exclusion. But naturally, being Nigeria, the lack of an officially-issued birth certificate is not necessarily an issue, as one contributor to the Nairaland Forum noted:

> *"Thats not a problem,its very easy to get it, xpecially the LG one. Just call a friend or family member in bauchi,tell him/her your local govment area,your date of birth, your papa and mama name. Thats all and he/she wuld go to your local govment 4 u and get it not more dan 2 days but* ***he gat to put sumtin in their hands*** *o. While the NPC can b done anywhere in 9ja but i havnt done mine o. It also* ***need puting something in their hands as well so as to be fast and removal of obstacles****. Na 9ja we dey o,nothing is imposible."*[65]

This situation is by no means limited to Nigeria; the authors have identified similar issues in Malawi and Liberia, and there is every reason to assume similar issues will be found across sub-Saharan Africa. In particular, the very low number of official registration centres in many of these countries is a significant problem.

But the problem in many countries, particularly in Africa and South Asia, is not necessarily limited to registration difficulties. There are often issues with the accuracy of information held on

[64] http://blogs.lse.ac.uk/humanrights/2015/05/28/questions-of-legal-identity-in-the-post-2015-development-agenda/
[65] From Nairaland.com: 'Nigerian Birth Certificate Needed'

birth certificates, due to the prevalence of multiple languages and the representation of those languages in written form. For example in Kenya, which has a relatively well organised system in which medical facilities provide a birth notification card, followed by birth registration at the local town hall (other arrangements are in place for births that take place outside medical facilities), there are occasions when discrepancies occur, so that a parent's identity card has a different spelling of the parent's family name from the child's family name on the birth certificate. This can give rise to problems in later life with the claiming of inheritances from parents, the issuance of identity and voters' cards, and of course passports and international travel. Even data such as birth dates and marriage certificates may be problematic and many registries have special investigation teams for issues like allegations of bigamy, proposed corrections to the records (including revised paternity details) etc.

Across the world, there are potential issues with birth certificates when a child is adopted – at this time, a child will commonly be issued with an updated birth certificate, with the original only being available on application to the proper authorities. This is further complicated in countries with a high prevalence of child abandonment, since the identity (and nationality) of foundlings can be a contentious issue. A prominent example is the Philippines, where a foundling has no automatic right to a birth certificate or nationality[66] despite Article 7 of Convention on the Rights of the Child[67]. It is the custom for their new parents to register their birth as though they are the biological parents – a so called 'simulated birth', but this can cause legal problems for the 'parents', and result in the child being denied nationality and a legal identity. The Philippines Government is seeking to address this by encouraging such 'parents' to legally adopt the child, so that they have a right to a legal Philippines identity through their adoptive parents, rather than in their own right[68].

#### *4.3.1.2 Other Documents*

Where a country has a poor track record of issuing birth certificates, this will commonly cause problems for citizens in later life, and this usually becomes apparent in contacts with authorities; for example, in accessing education, or trying to access some other form of social benefit, exercising the right to vote, receiving disbursements from international aid agencies or inheriting property. In these cases, a number of workarounds are commonly adopted, such as the acceptance of a letter from a local community leader or pastor 'sponsoring' the applicant (sometimes with a small payment to ease the process). However, this is obviously open to abuse – since the purpose of registration is generally to access some form of social benefit, it is common practice for such letters to be reused many times, and their value, and consequently the value of any identity documents that rely on them, is compromised.

This approach is generally used in the issuance of documents of lesser significance than an identity card; for example, voters' cards or driving licences. However, it is important to be aware of the inverted 'pyramid of trust' that is being created in these circumstances; at the root is the dubious letter from a pastor; built on that are voters' cards and driving licences; and built on them in turn are identity cards, bank accounts, passports etc[69], all potentially with no real link to a substantive identity.

[66] http://www.philstar.com/opinion/2015/10/01/1505841/no-international-law-confers-philippine-nationality-foundling
[67] http://www.unicef.org/crc/files/Rights_overview.pdf
[68] http://www.gmanetwork.com/news/story/501342/news/specialreports/questions-of-identity-foundlings-the-legally-adopted-and-their-political-rights
[69] Sadiq, K. 2009. Paper citizens: How illegal immigrants acquire citizenship in developing countries, Oxford: Oxford University Press.

### 4.3.2 Identity Cards

Like birth certificates, the issuance of identity cards in most countries in the developed world is a well established process. It is typically initiated by the applicant, who has a need for the identity card in order to assert their right to access services (education, employment etc).

Until relatively recently, the process followed a standard path. The applicant was required to present a fundamental identity document (usually the birth certificate), which was backed by the endorsement of a trusted member of society – traditionally a lawyer or teacher who has known the person for at least 5 years, who states that the applicant is indeed who he claims to be, and usually endorses a recent photograph. In recent years however this process has been substantially enhanced, to include a range of electronic checks. Indeed it is now the norm to rely heavily on checks with online credit reference agencies, who are able to verify that the applicant appears to be the person named in supporting documents (based on a history of financial transactions), that the claimed address is correct, and that this evidence has weight, since the records cover a reasonable period of time. Online comparisons with passport issuing agencies are also carried out, these having the benefit of having a photographic record of the applicant for comparison.

For those countries with compulsory national service, particularly amongst the emerging economies, it is commonplace to accept military ID documents when registering. This is assumed to be a reasonable approach, based on the view that someone who wasn't a citizen wouldn't willingly serve in a nation's armed services although age related requirements for military service might result in inaccurate birth dates being recorded.

At the point of issuance, additional details about the applicant are captured. For a digital identity, these are backed by some combination of biometrics[70] (iris, face, fingerprint etc). At the end of the process, a card is issued, with the card for a digital identity often being a smartcard of some form, carrying as a minimum a digital identity app capable of verifying the biometric profile of the bearer and providing assurance of identity, both locally (at a terminal) and remotely (via mobile phone or Internet).

At its core, the same registration and issuance process applies across the emerging economies, though the particular circumstances give rise to some substantial issues:

- In some countries, the issuance of an identity card is seen as an overtly political act, since it carries with it an automatic entitlement to vote. For example in Kenya, where Government has long been the purview of the dominant Kikuyu tribe, it has been common practice to put obstacles in the way of the issuance of identity cards to members of other tribes, such as Nubians.
- Multiple, duplicate applications for an identity card (arising for example from multiple users of a single sponsorship letter) can be avoided to a great degree by an online check with a national identity database. So it is the case, for example in Nigeria, that registration centres require broadband Internet connectivity. This is not something that

[70] There is an overwhelming tendency in the market to regard biometrics as close to infallible. However, Consult Hyperion's experience with biometrics, particularly in emerging economies, indicates that it is a technology to be used with caution, with strict attention to the characteristics of the citizens it is intended to serve and careful control of the registration process. Unfortunately, since this experience was gathered whilst working on client projects, no publicly available papers have been published.

is widely available in rural Nigeria, and this has led to significant problems with the rollout of the NIMC digital identity card.

- A further complication with digital identity rollouts in emerging economies is the lack of secure data centres to hold citizen data. Not unreasonably, local identity agencies typically require that citizens' data is held in-country, a difficult requirement to satisfy in a country with no server-grade data centres, intermittent power, and poor telecommunications infrastructure. These issues are discussed in more detail in Chapter 11.

### 4.3.3 Digital identity without identity cards

In countries such as the UK, US and Canada, digital identity schemes are being established that do not link to a national identity card (or associated register). These countries for historical, cultural and political reasons do not have a central identity register. The identification of individuals at the point of registration is done through a combination of sources that together corroborate and provide sufficient "assurance" in the digital identity being established. The UK's Verify programme requires evidence in three categories:

- "Citizen": Government issued official documents such as passports and driving licences.
- "Money": Credit reference data or other reliable financial evidence.
- "Living": Other sources such as utility bills that link the individual to the claimed address.

As well as positive evidence of identity, identity providers are required to check negative sources (e.g. counter fraud checks) to detect identities that may have been compromised.

The UK scheme allows a wide and growing range of documents to be provided as proof of identity. There is still, however, a significant number of people who are not able to establish digital identity. This could be because they do not have necessary breeder documents or have only been in the UK for a short period of time. For these costly face-to-face checks may be required.

## 4.4 Scope of digital identity schemes

Digital identity schemes tend to focus in one of two areas:

- **Logon**: this is the focus on many internet-based identity initiatives.

  OpenID Connect is entirely focused on federated logon to online services. The identity services offered by the big internet players, such as Facebook and Google, build on the same ideas. Their objectives are two-fold:

  - Increase engagement with customers
  - Use logon as a hook to acquire more data

  The UK government's Verify programme is also focused on logon as a means to enable their "Digital by Default" agenda as the base identity provided at logon is then matched to existing data held by the service provider. Whilst the motivation may differ from social networks the resultant digital identity services bear many similarities.

- **Digital signatures**: this has been the focus on many state-led eID programmes. This is particularly the case in Europe where many eID programmes were created as a response to the digital signature directive which defined high grade digital signature standards for interactions with government. Thus the eID smart card will explicitly include digital signature capabilities.

## 4.5 Digital identity information assets

In assessing the privacy impacts of identity programmes it will be necessary to consider the information assets that are collected, stored, processed and shared. The following table illustrates the range of information that may be associated with a digital identity system.

| Asset Type | Examples |
|---|---|
| Foundational register identifiers | National ID number |
| Functional register identifiers | Social Security Number |
| Core identity attributes | Name, Address, Date of Birth, Gender |
| Family information | Next of kin, family members |
| Ethnic information | Race, religion, place of birth, affiliation |
| Entitlements | Benefits, Vouchers |
| Stored value | Monetary value (prepaid) |
| Health information | Conditions, Allergies, Donor card, Disabilities |
| Transit | Tickets, Permits |
| Usage data | Transaction history, Service usage |
| Digital identifiers | Device identifiers, public key data |
| Biometric data | Fingerprints, Iris scans, Facial images |

*Table 2: Digital Identity Assets*

The range of information processed varies considerably between digital identity schemes. The Canadian credential brokerage service, for example, only processes digital identifiers. In the Nigerian digital identity scheme, the smart card will potentially contain applets (and data) relating to a range of government services.[71]

The location in which information is stored varies considerably as well. Databases may be operated centrally or spread across a distributed set of organisations. Information may also be replicated in the citizen's device (often a smart card).

The type of information, its storage and processing will all directly impact the privacy risks associated with a particular approach, as explored in Section 6.

[71] http://www.nimc.gov.ng/sites/default/files/id_card_policy.pdf, section 3.5.3

## 4.6 National eID schemes

As described above there are numerous national eID schemes (many of them state-led). Section 5.2 and Appendix A provide more information on a selection of these schemes examined during the course of preparing this report.

The following table highlights national eID schemes that have notable distinctive characteristics.

| Country | Distinctive feature(s) of eID scheme |
|---|---|
| Austria | • Uses unlinkable sector specific identifiers (and associated cryptographic keys and digital certificates) for privacy reasons.<br>• eID scheme able to link identifiers but clear boundaries around scope and purpose of that organisation. |
| Estonia | • Most mature eID scheme in world.<br>• Includes mobile version<br>• Focus on establishment of basic identity. Service providers (including government departments) responsible for establishing and maintaining service provider specific data. |
| Norway | • Operated by the banks.<br>• Follows a 4-party model, similar to card schemes and IdenTrust. Citizens are issued with digital identities by their banks. The scheme provides an overarching certificate authority (operated by NETS) allowing the digital assertions to be verified. |
| Pakistan | • Operated by autonomous entity NADRA.<br>• Used as a template for both Nigeria and Kenya eIDs |

*Table 3: Distinctive National eID Schemes*

## 4.7 Legal basis of digital identity

Any digital identity (e-identity) service is built on relationships between the identity provider, those whose digital identity it manages (individuals), and the service providers (government departments, private sector entities, etc). For such a service to be successful – that is, it is of value to the citizens and relying parties, and is sustainable over the long term – it is essential that all parties trust the service operator: citizens trust it to handle their private data securely and to manage their digital identity, and service providers trust it to ensure that the asserted identities are reliable (cost effectiveness is also a significant issue, but not the subject of this document).

Typically this multi-dimensional network of trust is built upon an Identity Trust Framework, the governance structure for a specific identity system that consists of[72]:

- The **technical and operational specifications** for the system, which define how the system works, the roles and responsibilities of participants, and provide adequate assurance that it is trustworthy;
- The **legal rules** that govern the identity system, and that regulate the content of the operational and technical specifications, make those specifications legally binding and enforceable, and define the legal rights, responsibilities and liabilities of the participants.

The legal rules are a combination of existing public law (including both specific identity-specific law, if it exists, and generally applicable laws around privacy and data protection, warranty, e-transactions, etc.), supplemented by contractual agreements between the parties and any standards adopted by the parties (usually specified in the contract).

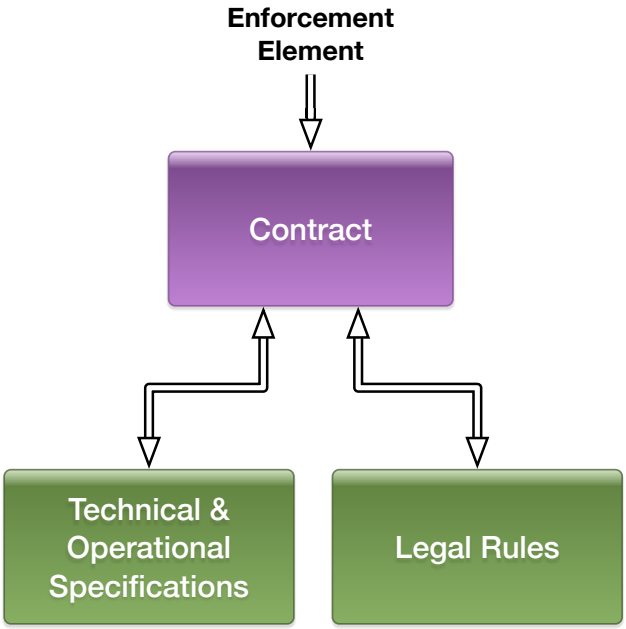



*Figure 7: Legal Basis of a Trust Framework*

The legal rules that an identity trust framework must comply with are naturally country-specific, but there are some general comments that can be made. Leaving aside regulations around privacy and data protection, which are dealt with elsewhere in this document, there are many regulatory areas that touch on legal identities[73] and the establishment and operation of an identity trust framework:

- A vital element to the legal rules is legislation giving electronic signatures the same legal standing as handwritten signatures. In the EU, this came with the 1999 E-Signatures Directive.

[72] ABA's proposed definition of an identity trust framework: apps.americanbar.org/dch/thedl.cfm?filename=/CL320041/newsletterpubs/4-Trust-Framework-and-Liability-Overview.ppt

[73] https://www.nesta.org.uk/sites/default/files/research_on_digital_identity_ecosystems.pdf

- Competition law, which is chiefly concerned with creating efficient markets that give citizens choice. Clearly a single national identity trust framework could potentially fall foul of competition law – whether or not this is a substantive issue for a government-run national identity scheme is a moot point. Instead, it is principally an issue for private-sector identity schemes.

  A key aspect of competition law is market dominance, but this only becomes troublesome when it is abused to unfairly exclude competitors, or exploited in a way that harms consumers. One aspect that arises from this concern is that of **interoperability**; so identity trust frameworks should seek to become interoperable with others in their sector or country of operation. This is dealt with later in this document.

  Issues around mergers of identity trust frameworks can also give rise to concern amongst regulators, due to the potential to reach market dominance without due consideration of the issues arising.

- Consumer protection, which creates an imperative to protect citizens from risks arising from the use of an identity trust framework, specifically potential harms caused by abuses of personal information or service outages.

- The EU has probably progressed further in e-identity legislation than any other legislative body. There have been a number of legislative initiatives that impact on identity trust frameworks:

  - Most relevant of all, the European Directive on electronic identification and trust services for electronic transactions in the EU (the **eIDAS** Regulation) of July 2014. This Directive doesn't make digital identity (eID) mandatory, but does aim to greatly increase the mutual recognition of eID between countries, in order to facilitate cross-border business as well as international administrative tasks for citizens. To this end it ensures that people and businesses can use their own national electronic identification schemes (eIDs) to access public services in other EU countries ***where eIDs are available***, and creates a new EU market for "electronic Trust Services" (eTS) - namely electronic signatures, electronic seals, time stamps, electronic delivery services and website authentication - by ensuring that they will work across borders and have the same legal status as traditional paper based processes.

  - The Directive on the Reuse of Public Sector Information is intended to promote open data, which is likely to result in public registers containing personal information becoming open data, and so allowing organisations to build profiles of EU citizens.

  - The Digital Agenda is aimed at the creation of a Digital Single Market across the EU. This implies a requirement for interoperability for the implementation of e-government, through the use of compatible standards and technical specifications.

Interesting as they are, the EU initiatives are of course limited in their scope to the EU. However, it is notable that the European model of data protection has already been adopted by other countries around the world, such as the Philippines. Privacy advocates tend to favour the

EU model because it is strong and tested, and local policymakers also wish to have their country given special status for data transferred from the EU. It is therefore worth considering that other EU legislative initiatives, in particular eIDAS, may influence regulation elsewhere in the world over the coming years.

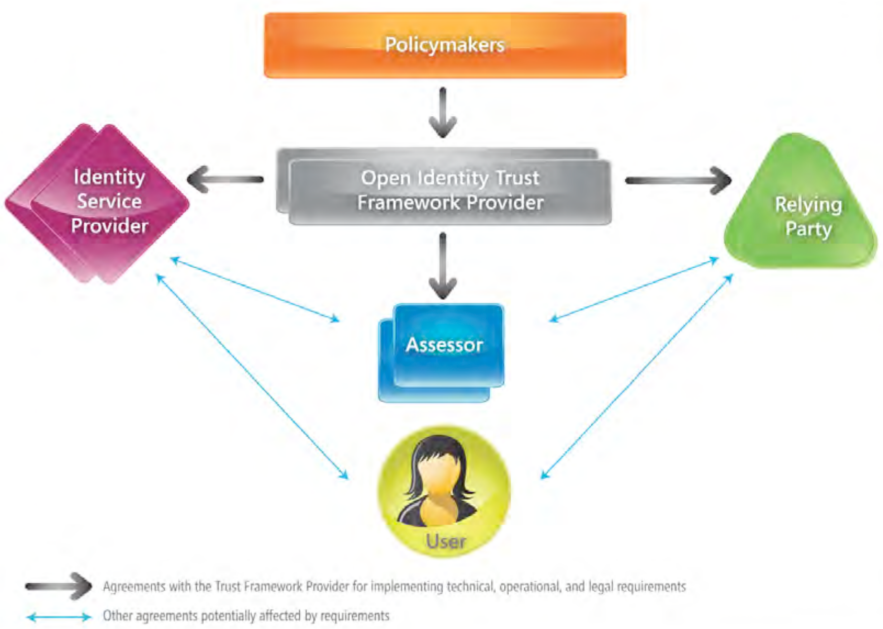



*Figure 8: Open Identity Trust Framework (OITF)*

All digital identity services must inhabit this legislative environment – it is certainly not limited to governmental initiatives, and applies equally to private sector initiatives such as those inspired by the Open Identity Exchange (OIX) and their Open Identity Trust Framework (OITF). This can be seen in OIX's own description of the working of OITF:

> *Policymakers start by deciding the technical, operational, and legal requirements for exchanges of identity information that fall under their authority. They then select OITF Providers to implement these requirements. These OITF Providers translate the requirements into a blueprint for a trust framework that may include additional conditions of the OITF Provider. The OITF Provider vets identity service providers and relying parties and contracts with them to follow its trust framework requirements when conducting exchanges of identity information. The contracts carry provisions relating to dispute resolvers and auditors for contract interpretation and enforcement. Requirements flow down through agreements, as shown in the directional arrows (in* Figure 8*)*

## 4.8 Why digital identity schemes fail

Digital identity is not a new topic. There have been numerous attempts to establish digital identity schemes in the past. Some, such as Identrust[74], have been successful in finding niche markets. Others, such as Microsoft's Infocard[75], have been abandoned along the way. There are several factors that have hindered the growth of digital identity in the past:

- **Smart card based eID schemes have been designed for niche uses**. A key driver for many eID cards (especially in Europe) has been the digital signature legislation. As a consequence these schemes have tended to focus on low volume high value applications such as contract signing. Furthermore smartcards are not well suited to online services due to the need for readers which are not ubiquitous. There have been efforts to integrate contactless smart card reader technology into standards for personal computing devices[76] but this has not become widespread in use, probably due to the focus shifting to mobile devices.
- **Commercial models have often been the stumbling block for identity initiatives**. Whilst enterprises have been used to paying handsomely for VPN hardware tokens (e.g. RSA SecurID), consumers have not. Furthermore, identity transactions do not involve a direct flow of money and so, unlike payments, there is no opportunity to take a slice of the transaction value.
- **Obtaining critical mass is always a challenge**. There is a chicken and egg problem. As a consumer, an identity scheme is only of interest if it is accepted by a wide variety of service providers. As a service provider an identity scheme is only of interest if it is used by a significant proportion of the service provider's customers. Therefore, both conditions need to be satisfied simultaneously.
- **Liability has often been a stumbling block in the past**. Unlike payments, the potential losses could be much greater than the value of the service being used (e.g. if incorrect identification leads to injury or loss).
- **The user experience offered by identity services often was not optimal**. This was either because the services were not completely intuitive or because they required additional hardware (e.g. smart card reader) that the user needed to remember to carry with them.

By way of contrast, the internet giants (Facebook, Google, Apple etc) have amassed billions of customers and established ubiquitous logon methods. These federated logon[77] methods enable the internet giants to collect data from third parties to add to the huge data assets they already manage.

[74] https://www.identrust.com/company/index.html
[75] https://msdn.microsoft.com/en-us/magazine/cc163626.aspx
[76] For example, Intel's IPT technology
[77] http://en.wikipedia.org/wiki/Federated_identity

# 5 PRIVACY OF DIGITAL IDENTITY IN PRACTICE

Assessing the impacts of digital identity schemes on the privacy of individuals is overwhelmingly a contextual exercise, requiring a consideration of the legal, societal, cultural, technological and security factors relevant to each unique digital identity system. A government eID system in one country may have a different impact on privacy than the same system in another country, depending on the design of the system and the context in which it is operated.

Nevertheless, it is possible to derive some general conclusions about the types of impacts that could be expected to flow from the deployment of particular digital identity schemes. These impacts can be viewed from three perspectives:

- Regulatory: to what extent does the digital identity system incorporate accountability, transparency and strong governance mechanisms to guard against adverse impacts on individuals' privacy?
- Technology: how do particularly technological choices impact upon individuals' privacy, security and effective utilisation of the system?
- Commercial: how does the digital identity scheme work commercially? Does the commercial model encourage or discourage protecting the privacy of individuals? Does the commercial model represent a "fair deal" for the citizen? Is the scheme commercially viable, given this will be key to its long term success, particularly if public funding is limited?

Viewing a digital identity scheme from each of these three perspectives enables a holistic understanding of the trade-offs inherent in various schemes.

## 5.1 Perspectives

### 5.1.1 Regulatory perspective

The extent to which a particular digital identity scheme provides for strong governance and accountability mechanisms conducive to enhancing privacy will depend, to a large degree, on the regulatory environment in which that scheme is being deployed, as well as the legacy systems in place when a new scheme is being designed and implemented. In this regard, the following questions, among others, will be relevant:

- **Is there a comprehensive data protection law pertaining to both private and public entities in the country in which the digital identity scheme is operating?**

  Some countries only regulate the collection of data by either private or public entities, leaving gaps for example when unregulated government service providers process personal information, or alternatively when private sector entities providing identification or authentication services do the same.

- **Are all private sector identity providers and credentialed service providers processing data within a jurisdiction where a comprehensive data protection law exists?**

  Where identity providers have foreign subsidiaries or parent companies, or utilise overseas cloud storage services, data collected in one jurisdiction may be transferred to another for processing. It is essential that any additional jurisdiction in which data is processed has equivalent data protection standards to that in which the data is collected, and that contractual conditions to this end are imposed during the procurement process and this is not covered by national law.

- **Is there a competent independent authority charged with monitoring compliance with and remedying contravention of data protection laws within the country?**

  Without an effective enforcement mechanism in place to remedy data breaches and penalise non-compliant data processors, data protection legislation may be ineffective.

- **How are the relevant terms, such as personal data, sensitive data, and data processing defined under the applicable legal regime?**

  Analysing different approaches to defining personal data, in particular, provides an important tool for understanding how a country's regulatory system will impact uniquely on the level of protections for privacy provided by an individual digital identity scheme. There is variance, even amongst countries that follow the EU model, in how these terms are defined. For example:

  - Generally, *personal data* will be defined as "any information relating to an identified or identifiable data subject." The interpretation of "identifiable" changes by country; in some, it means "identifiable by virtue of other information in the possession of the data processor"; in others, it means "identifiably by virtue of other information reasonably obtainable by the data processor" – a quite different approach. Further, anonymised data is generally not treated as personal data, whereas pseudonymised data is.

    Two pertinent examples: Malaysian law provides that personal data is defined as data *in respect of commercial transactions* only; and European law generally provides that data protection obligations are owed only in respect of the personal data of living persons.

  - *Sensitive data* is defined only by the categories to which it is applied, such as health information, political membership information, religious affiliation. Some countries, such as Estonia, Peru and the UK, prescribe that sensitive data include a broader range of data, including membership of a trade union and criminal records.

  - *Data processing* is a term incapable of exhaustive definition. Generally, it will refer to collection, storage, and handling in any manner. Recently, the European Court of Justice found that the act of returning search results in response to a query based on an individual's name was an act of data processing with respect to which Google was required to abide by European Data Protection law. This will continue to be a growing area, especially with the

advent of the Internet of Things: consider, if a lightbulb is able to transmit and receive data such as IP addresses, will that constitute an act of data processing?

- **Does the country have strong legal and institutional mechanisms to bolster data protection laws?**

  Laws which establish rights and responsibilities may be meaningless without strong and independent courts and well-resourced public authorities to ensure compliance with those laws.

- **How have courts and other authorities interpreted those provisions with respect to particular types of data or processing**?

  For example, under the EU Data Protection Directive, personal data is defined as information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.[78] Yet different interpretations of that definition have resulted in EU member states taking different stances with respect to particular pieces of data, including static and dynamic IP addresses, and anonymized data.

The table below illustrates that, even within the EU, member states have different interpretations of what constitutes personal data:

| **Differences in EU member states' interpretation of personal data**[79] | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Data** | **CZE** | **FRA** | **DEU** | **LUX** | **ESP** | **SWE** | **GBR** |
| A static IP address (other than in the hands of the subscriber's ISP) | ✓ | × | ✓ | ✓ | ✓ | ✓ | × |
| A dynamic IP address (other than in the hands of the subscriber's ISP) | ✓ | × | ? | ✓ | ✓ | ✓ | × |
| A cookie | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Data that is coded[80], when the person has the keys to the code | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Data that is coded, when the person does not have the keys to the code | ✓ | ? | × | × | × | ✓ | × |

*Table 4: Differing EU State Attitudes to Personal Data*

[78] For further exploration of the full application of the definition of personal data, see Article 29 Data Protection Working Party, *Opinion 4/2007 on the concept of personal data*, 01248/07/EN WP 136, adopted 20 June 2007, available at http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf.

[79] Richard Cumbley and Peter Church, "EU – What is Personal Data?", *Linklaters: Technology, Media and Telecommunications News*, 1 October 2008, available at http://www.linklaters.com/Insights/Publication1403Newsletter/PublicationIssue20081001/Pages/PublicationIssueItem3513.aspx.

[80] Data that has been coded so that direct identifiers of a person (like names and addresses) are replaced by a key or pseudonym, and can be linked to an identifiable individual only by a limited number of parties.

### 5.1.2 Indicators of strength of privacy governance

It is difficult to generalise about the factors that contribute to a "strong" data protection environment in any particular country. However, we believe that the two preeminent indicators of the strength of privacy governance in a particular country are:

- the existence of a comprehensive data protection law, and
- the existence of a specialised, independent data protection authority that is adequately funded[81].

Below we illustrate the strength of the data protection regimes in the countries in focus in this report, based on these indicators.

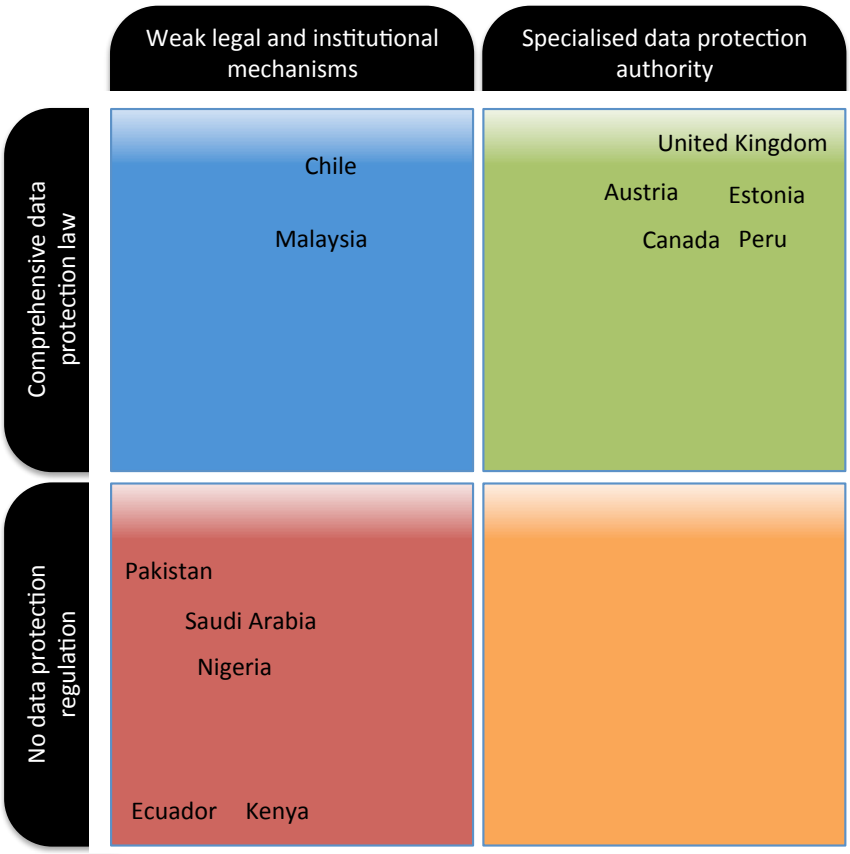



*Figure 9: Plot of Countries' Data Protection Regimes*

[81] Greenleaf, G. 2012. "Independence of data privacy authorities (Part I): International standards," Computer Law & Security Review (28:1), pp. 3–13 (doi: 10.1016/j.clsr.2011.12.001) provides guidance on how to assess the independence of data privacy authorities.

### 5.1.3 Technology perspective

There are several technologies employed in the delivery of digital identity systems as illustrated in the following diagram:

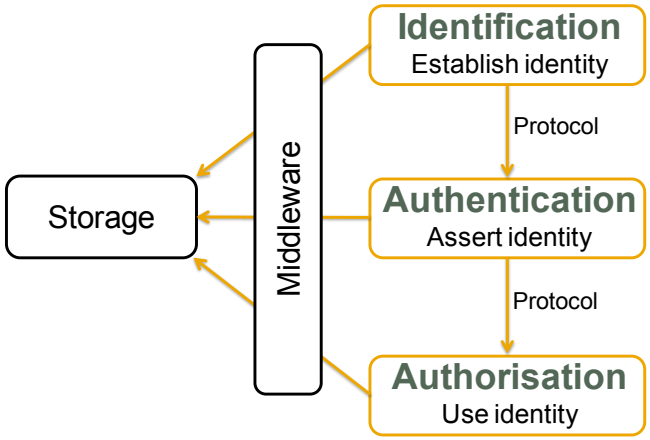



*Figure 10, Identity Technologies*

These can be put together in a variety of ways to deliver an end-to-end system. The categories we consider here are:

- **Identification**: Technologies that assist in establishing the identity of an individual.
- **Authentication**: Technologies that enable an individual to assert an established identity. Typically this will be through the use of a token or biometric.
- **Authorisation**: Technologies to provide users with the means to control access to their data and resources.
- **Storage**: The means by which verified identity attributes are stored.
- **Middleware**: Integration technology allowing service providers to gain access to verified identity attributes held in the storage. Middleware can be local (e.g. on a PC enabling integration to a smart card) or remote (e.g. providing integration to online identity services).
- **Protocol**: The mechanism for performing digital identity transactions including defining the sequence of interactions and data flows between the multiple parties involved in an identity transaction.

Table 5 below considers different technologies under these categories, and assigns a privacy score which is intended to indicate how well the technology supports the protection of personal information relative to the other technologies.

| Technology | Type | Privacy Score | Rationale |
|---|---|---|---|
| EMV[82] | Authentication | H | EMV uses strong cryptographic security. This can be provided via a smart card, a secure element within a mobile phone or potentially via a "hardened" app in a mobile phone.[83] |
| SIM[84] | Authentication | H | A GSM-compliant mobile phone's SIM is a specialised smart card (qv), and offers a tamper resistant cryptographic environment. It can host apps created in the SIM Toolkit (STK) environment, which can use encryption for transactions and general communication. It can support communication over any of the mobile phone's network connections, including mobile data, SMS and USSD. |
| Smart Card[85] | Authentication | H | Tamper resistance cryptographic hardware. Established and recognised secure technology. |
| Physiological Biometric[86] | Identification or Authentication | H | For identification needs, multiple biometrics are necessary to establish uniqueness. By contrast, for authentication a single biometric can provide effective means of authentication of asserted identity. |
| Behavioural Biometric[87] | Authentication | M | The technology is less mature than physiological biometrics (qv), and more aligned to risk management than absolute or explicit authentication. |
| Mobile App | Authentication | M | Can be high security, depending on protection built into the app. |
| Risk based authentication (RBA)[88] | Authentication | M | Like behavioural biometrics, RBA provides corroborating evidence for authentication rather than explicit authentication. No standard way to measure performance. |

[82] EMV (Europay, MasterCard, Visa) is the set of standards for worldwide interoperability and acceptance of secure payment transactions.
[83] HCE (Host Card Emulation) has caused the payments industry to consider software approaches to EMV leveraging techniques such as white box cryptography. In these solutions, authentication credentials are usually tokenised to reduce the risk associated with the compromise of a credential. Typically the numerous measures including software hardening and server side risk monitoring are employed to ensure the overall residual risk is acceptable.
[84] Subscriber Identity Module
[85] A smart card is a device that includes an embedded integrated circuit that can be either a secure microcontroller with internal memory or a memory chip alone. The card connects to a reader with either direct physical contact or with a remote contactless radio frequency interface.
[86] Physiological biometrics is the field of study related to the measurement of innate human characteristics such as fingerprints or iris patterns.
[87] Behavioral biometrics is the field of study related to the measure of uniquely identifying and measurable patterns in human activities, rather than innate human characteristics.
[88] Risk-based authentication is a dynamic authentication system which takes into account the profile of the agent requesting access to the system or service in order to determine the risk profile associated with that transaction. The risk profile is then used to determine the complexity of the challenge required.

| Technology | Type | Privacy Score | Rationale |
|---|---|---|---|
| SMS | Authentication | M | SMS used straight relies on mobile network encryption, which is known to be weak. It needs to be augmented with application security, implemented for example in a mobile app or in a SIM Toolkit app (see SIM), both of which can use SMS as the bearer technology. |
| TAN List[89] | Authentication | L | Access to physical list required but can be easily copied once access is obtained. |
| OAuth | Authorisation | M | A protocol for providing access tokens (which may be temporary) to allow third party applications to access resources (data) on behalf of the resource (data) owner. |
| UMA (User Managed Access) | Authorisation | M | A recently established standard that defines how a resource owner (e.g. an individual) can control access to their resources (e.g. personal data) by third parties. The standard was developed by the Kantara Initiative[90]. It builds on and extends OAuth (qv). |
| Scanning documents | Identification | H | Digital validation of documents using image processing; a relatively new technology but thought to be robust. AU10TIX is a leading vendor in this space[91]. |
| Credit Reference Agency Data | Identification | M | In developed markets this is a de facto method of establishing identity. Can appear invasive, where knowledge based question is generated from credit data. |
| Government Registries | Identification | M | Usually viewed as authoritative. Anecdotally can often contain significant numbers of fraudulent identities. Privacy will depend on amount of data held and control of access to it. |
| Social Identity Verification | Identification | L | The use of social media data including self asserted data and social graph to establish identity. The strength of such an approach is unproven. It relies on sufficient data being readable. |
| Agent | Middleware | H | A decentralised broker for connecting parties together in an identity transaction. Typically the agent will run on a client device. It will still need to access centralised directories for the purposes of discovery but this approach prevents data aggregation. |

[89] Transactional Access Numbers (TAN) list – a printed list of codes from which the user is asked to select one, as a means of authentication. Used in Danish eID system NemID (https://www.nemid.nu/dk-da/om_nemid/sikkerhed/teknikken_bag_nemid/).
[90] http://kantarainitiative.org/
[91] http://www.au10tix.com/index.php/products/front-end-solutions/

| Technology | Type | Privacy Score | Rationale |
|---|---|---|---|
| Integration Layer | Middleware | L | A general layer or infrastructure that service providers plug into for accessing digital identity services. This component is non-specific but can be seen, for example as component of both the World Bank's ID4D model for digital identity and the Estonia digital identity infrastructure. Similar to a hub, it presents aggregation risks. |
| Hub | Middleware | M | A centralised broker for connecting parties together in an identity transaction. An alternative to the more distributed federation offered by OAuth. Hub and spoke makes session security easier to achieve however the hub is a point of aggregation and so presents a privacy risk. |
| OpenID Connect | Protocol | M | Lightweight internet protocol built on OAuth[92] for performing federated logon. To be secure requires strong mutual authentication to be provided (outside of the protocol) |
| SAML (Security Assertion Markup Language)[93] | Protocol | M | SAML is robust, but relies for security on correct implementation of appropriate PKI[94] outside of the protocol. |
| Personal data store (PDS)[95] | Storage | H | Each PDS (or the data within it) will be encrypted by a key under the end user's control. Hence a data breach which does not expose the end user's key would not reveal any personal data. |
| Cloud databases[96] | Storage | L | Cloud databases protected only by online authentication are susceptible to data breaches. |

*Table 5: Identity Technologies*

Technological choices will impact upon the privacy afforded to the individual in any digital identity system. The following sections discuss the technology types in the table above, highlighting where there are technologies that, when used well, enhance privacy and those that have the potential to negatively impact privacy.

#### *5.1.3.1 Identification Technology*

This is an area of significant innovation being driven by a desire to make services fully digital and remove friction wherever possible. In particular, technologies are being developed (such as analysing photographs of identity documents captured with the user's mobile phone) to allow

[92] Internet protocol that creates tokens for access to resources, access can be bound by scope or time.
[93] SAML is an XML standard that allows secure web domains to exchange user authentication and authorization data. Using SAML, an online service provider can contact a separate online identity provider to authenticate users who are trying to access secure content.
[94] Public Key Infrastructure, the management of cryptographic keys and certificates that underpins internet security and digital identity.
[95] A PDS is a service which enables an individual to store, manage and deploy their key personal data in a highly secure and structured way. It gives the user a central point of control for their personal information. A PDS might be located on a user's PC, but may also be held in the cloud.
[96] Databases running on a cloud computing platform.

immediate enrolment into services, avoiding the need for a manual step (such as attending an office) that breaks the process.

The use of social graph data (for example, as offered by Trulioo or Veridu) to complement traditional sources of digital identity has the potential to provide new means of identification, even though it uses data that is self-asserted. The theory is that, with enough data and connections between users (i.e. social graph), an individual can be identified to a high level of assurance. This may provide a means to identify individuals who cannot be addressed through the traditional approaches to identification, due to being excluded. However, no metrics exist against which such processes can be measured, and from a privacy perspective there is a risk around the need for large quantities of data and the connections between individuals to be openly available for analysis.

The use of image processing to determine the genuineness of paper identity documents has the potential to be relatively robust especially where documents containing photographs are presented which can be compared against an image of the individual presenting the document (such as the services offered by Au10tix, Jumio and Mitek). To be even more robust, the ability to perform an online lookup to authoritative source (such as a passport database) would allow the validity of documents to be confirmed as well. The privacy aspects of this technology will depend very much on how images are secured during collection, transmission, analysis and archival, as well on what is done with those images. To be privacy respecting, the images should ideally be used for the sole purpose of identification.

Conventional databases such as government registries and credit reference agencies are the primary means of identification in the more developed countries. They contain large pools of valuable data and hence need appropriate protections including limitations on use. The use of “knowledge based authentication" to establish identity may use the data held by such organisations, especially credit reference agencies[97]. The citizen will be asked to answer a question pertaining to a transaction that is recorded in the database, to demonstrate knowledge that is only likely to belong to the citizen in question. Whilst this may not alter the overall risks to the data significantly (depending on how the service is implemented), it may appear sinister to the citizen. For example, in the UK’s Verify service when a citizen enrols with an identity provider they may be asked to confirm the balance of a recent credit card bill. This uses credit reference data that will not pass to the government, nor should it be stored by the identity provider, but this may not be obvious to the citizen.

#### *5.1.3.2 Authentication Technology*

The role of authentication technology in preserving privacy is primarily in ensuring that only the legitimate user can assert their digital identity.

Smart cards are widely recognised as a robust solution, being used in many eID systems and across the payments industry (especially card payments using the EMV[98] standard). They put the cryptographic keys used to protect identity into the hands of the citizen. However, the use of smart cards for accessing remote digital services has often been hampered by the need for peripheral reader devices, which introduce additional cost and inconvenience.

[97] http://lexisnexis.com/risk/downloads/idm/role-of-knowledge-based-authentication-in-identity-proofing.pdf
[98] Europay MasterCard Visa; see Table 5.

Mobile phones are becoming increasingly important for authentication, as they combine many security features with a user interface and connectivity, making the mobile device a versatile authentication device. They contain smart card technology (in the form of a SIM or Secure Element) which can provide strong protection of the citizen's cryptographic keys. This secure hardware is typically controlled by a third party (generally the mobile operator or the handset manufacturer) and so cannot be readily used without a commercial relationship with that third party. In this regard, the GSMA's Mobile Connect initiative is concerned with enabling digital use of MNO-controlled secure hardware (SIMs), and to be successful will need to include a realistic commercial model. Further, this MNO-control of the SIM has the potential to limit the use of SIM-based authentication, since the MNOs act as gatekeeper and control third party access.

It is of course commonplace for people in emerging economies – especially those at the Bottom of the Pyramid (BoP) – to share a mobile phone, either as a device shared amongst friends or family, or by renting one from a local entrepreneur. This does not necessarily diminish the potential for the mobile phone to be used for authentication, if the authentication is SIM-based, as even in cases where someone doesn't own a mobile phone they usually own their own SIM, and place it in the shared mobile phone before use. Of course, this does not apply to those cases where the authentication is app-based, on a smartphone for example.

The benefits of using the SIM for mobile phone-based authentication have recently been demonstrated by the "Error 53" issues that recently affected users of Apple's iPhones[99]. These devices use the "Touch ID" device to read the mobile phone user's fingerprints, and provide authentication by comparing them with the registered profile(s) stored in a secure enclave (analogous to the SIM, but controlled by Apple, not the MNO). Unfortunately, if the phone is repaired by unauthorised (by Apple) cheaper third parties, the integrity of the link between the Touch ID device and the secure enclave can be broken, so that the device cannot be relied upon for authenticating access to secure services (such as payments). The iPhone can detect this breakdown, but initially instead of blocking access to secure services, Apple chose to block all access to the mobile phone, a perceived over-reaction that was soon amended. The issue here is not Apple's over-reaction, but the vulnerability of a non-removable secure enclave to breakdowns of trust such as this.

Secure authentication services can increasingly be delivered over the top (without involving the mobile operator). The FIDO Alliance is focused on standardising authentication utilising authentication capabilities of mobile devices including biometrics. Since Google opened up the NFC interface in Android 4.4 (KitKat) the payments industry has been developing mobile payments solutions that are software based using a combination of software hardening techniques[100], tokenisation[101] and device management to provide a more dynamic approach to securing and managing the risks associated with payments credentials.

Biometric technologies (especially physiological biometrics) are especially interesting as they do not require the citizen to own or be issued with a card or device. It is important to distinguish between the two uses of biometrics. When used for *identification,* the biometric data captured needs to be sufficiently unique that the citizen can be distinguished from all other citizens in the population. For *authentication*, however, a much lower threshold of accuracy is likely to be

[99] http://www.theguardian.com/money/2016/feb/05/error-53-apple-iphone-software-update-handset-worthless-third-party-repair

[100] Techniques such as obfuscation, white box cryptographic, anti-tamper, anti-debug and jailbreak detection.

[101] Replacing a long lived high value credential with a short lived limited use credential which has lower risk allowing software based protections to be accepted as adequate.

acceptable, based on a 1:1 match with the (likely) owner of the device. The authentication merely needs to ensure with sufficient confidence that the presented biometric matches the identity being asserted. An analogy can be drawn with PIN entry. Most bank cards require a 4-digit PIN to be entered for authentication, meaning that there are only 10,000 possible permutations. Usually after 3 incorrect attempts the bank card is blocked. As an authentication mechanism this does not establish uniqueness (within the population there will be many people sharing the same PIN) but it is good enough to protect the card against misuse. It should be the case therefore that biometrics used for authentication (as opposed to identification) carry fewer privacy concerns, depending on the system design.

Privacy issues arise when authentication is either not performed or is carried out using weaker technology (e.g. passwords). These allow false identity claims, account takeover and leakage of data.

#### *5.1.3.3 Authorisation Technology*

Recently the focus of the internet identity standards community has shifted from the basic logon mechanisms (provided by OpenID Connect) to the more nuanced requirements of managing what data is shared and when. This has been a key focus of the Kantara Initiative, who published the "User Managed Access" standard in March 2015. There are some open source implementations but it is too soon for it to be widely adopted.

OAuth on the other hand, on which UMA is based, is widely used in internet applications. OAuth is less rich than UMA, which supports amongst other things the ability to centralise the authorisation management from the user's perspective and the ability for parties to share data (if an appropriate policy allows it) without the user needing to be present.

Providing effective management of data to users will ultimately be key to addressing the issue of consent.

#### *5.1.3.4 Protocols*

Protocols such as OpenID Connect are used to enable federated identity. They establish a mechanism under which an identity token can be passed from the identity provider of the citizen's choice to the service being accessed by the citizen. In OpenID Connect this involves a "bearer token" being created by the identity provider to be passed to the service provider. The issue with this is that presentation of the bearer token alone may be sufficient to access the service. In other words if a malicious application is able to capture the bearer token then it would be able to gain access to the service. This is recognised by the OpenID Foundation[102]. To avoid this issue additional layers of security are required, ideally mutually authenticated TLS[103], meaning that all parties are authenticated through the TLS to each other and in both directions. The vast majority of web services today only authenticate in one direction (web site to browser) which leaves open the possibility of man-in-the-middle and other attacks.

SAML is the other protocol often cited in the context of federated identity. It too has limited scope and is focused on standardising how security assertions can be packaged up into XML messages. The burden of security implementation including end-to-end authentication, securing sessions and so forth, is not addressed in this standard.

[102] https://openid.net/2015/05/26/enhancing-oauth-security-for-mobile-applications-with-pkse/
[103] Transport Layer Security, the cryptographic security used to protect the web site access and the successor to SSL (Secure Sockets Layer).

The technology to secure federated identity communications exists. Appropriate use of PKI and TLS in support of OpenID Connect and SAML, for example, should lead to a satisfactory solution. It is not enough however to simply claim that OpenID Connect or SAML is being used.

#### *5.1.3.5 Middleware*

"Middleware" is a generic term we use for the integration software that enables service providers to access and make use of digital identities.

In eID systems, client-side software is often deployed on the citizen's PC to provide an API for service provider applications to call, to access the citizen's eID. In the case of the Personal IDP model (see 6.1.6), the user agent software acts as a filter ensuring that data is only shared with the user's consent. Client-side software can be damaging to privacy. For both benign and malicious reasons client-side software can record usage data and send it to a third party. Where client-side software is verifiably designed for specific and legitimate digital identity purposes, this should result in a system that supports the privacy objectives. Deploying software onto untrusted devices (i.e. any device sourced by the citizen) will always present challenges and therefore careful consideration of the potential risks will be needed as well as ongoing monitoring of issues.

Server-side integration infrastructure including hubs and other integration services (such as the X-Road backbone in Estonia[104]) can be also be thought of as middleware. The risks here will be different, the primary privacy risk being that such services will provide a point of aggregation of digital identity data from all citizens using the eID service.

#### *5.1.3.6 Storage*

The final key component of a digital identity system is storage. There are two logical models for the storage of digital identity data.

Data can be stored in a centralised database. This is the traditional approach for all IT infrastructure including governments, banks, mobile operators and others. In this model the data is controlled by the identity provider and the citizen is dependent on the identity provider to put appropriate logical access controls around their data, but at some point it is likely that the data of many citizens is available in bulk.

Data can be stored in user controlled personal data stores. These could be located on a user device or (more likely) held in an online server. The crucial difference with centralised databases is that the data store for each citizen will be encrypted using a key which only the citizen has access to. This means that the identity provider themselves may have no access to the citizen's personal data. Consideration should be given to how to avoid loss of service if the citizen loses their key (or the personal device which holds their key), but in general this is an approach that supports good privacy goals.

### 5.1.4 Commercial perspective

It is important to consider the commercial model that underpins a digital identity scheme, since it is the viability of this model that will be key to its long term success if even a single element of the delivery of the scheme is reliant upon a private sector, profit-driven entity. In many cases, it is perhaps most appropriate to think of it as an identity infrastructure that is paid for centrally and provides services and benefits to all. Of particular interest in the context of this report is the

[104] https://e-estonia.com/component/x-road/

impact the commercial model has on questions of privacy; at the most simplistic, is it in the commercial interest of a participating organisation to compromise on individuals' privacy or is there an adequate alternative income stream to offset this risk? And does the model present a fair deal for the citizen?

There are a range of different commercial models for digital identity, ranging from the pure profit-driven to the government- or individual-centric. These are introduced below, together with high-level consideration of the privacy aspects of each.

- Those familiar models where citizen data is acquired by a commercial entity, and then analysed and sold directly or indirectly (targeted advertising).

  The principal users of this commercial model are the purely private sector IDPs, since this model clearly incentivises IDPs to aggregate and share as much data as possible. This clearly represents a threat to the registered citizens' privacy.

- Government pays for digital identity (some state-led eID schemes adopt this model).

  This is not a nakedly commercial model, though of course governments operating a digital identity scheme are required to operate within a budget to cover the cost of the service. Nonetheless, they are generally incentivised not to share citizens' data inappropriately, and any threat to the registered citizens' privacy is much diminished.

  However, this model can lead to a lack of clarity on who the 'customer' is (and indeed a view that there is no customer), which can cause a degree of complacency and allow poor practices to emerge, which can in themselves represent a passive threat to privacy.

- Citizen pays for digital identity as they pay for passports and driving licenses (some state led eID schemes adopt this model).

  As in the previous case, this is not a purely commercial model, and many of the same comments apply. However, there is an important distinction; since the citizen pays, there can be a much firmer understanding within the IDP of who the 'customer' is, which can result in an increased focus on the customer's needs – hopefully resulting in a more trenchant approach to privacy. In general, given the requirements for states to issue basic identity credentials for free, this implies that digital identities might be seen as an additional, rather than default, service offering.

- Service provider (consumer of digital identity) pays for digital identity (for example, UK Verify and the Canadian model; this is also the usual model for access to credit reference agency data). There are two variants of this model:

  - Service provider pays centralised or brokered identity provider (who manages identity on behalf of citizen): Overall, this model has something in common with the first, purely commercial model, the principal driver for difference being that payment is by the service provider, rather than arising through advertising. So, for example, the IDP might be incentivised to analyse citizen data, drawing conclusions that allow it to better market its services to service providers. This can be expected to have some impact on citizen privacy, but this is likely to be minimal.

- Service provider pays personal data store provider and citizen, who manages own identity via data store: This introduces the concept that some of the funds may flow back to the citizen, paying them for access to data. To date such models have focused on providing citizens with an "incentive" payment rather than something commensurate with the full value of the data being shared. The operators of personal data stores still need to be paid.

Other commercial models exist in the digital identity market, one example being a model based on a revenue share of a digital identity-enabled business, but as these do not directly translate to foundational digital identity systems, they are not considered further.

At first sight, it might appear from this high level analysis that the 'citizen pays' model would be preferred over the 'government pays' model, but of course things are never that simple. Few citizens are willing to pay for an identity card, or the associated access to state services that require a digital identity, since they see it as the state's responsibility and certainly don't see the value to themselves. This leads identity providers to one of the other commercial models in the large majority of cases.

## 5.2 Examples

The following examples outline how in a selection of countries the regulatory, technical and commercial aspects of the national digital identity scheme(s) are addressed, highlighting the positives and negatives of each, from a privacy perspective.

From these examples, it can be seen that there is often a flow from regulation into a scheme's technical and commercial features. In particular, where a country has a good privacy regime this often results in a stronger focus on privacy in implementation. However, the link between regulation and the commercial characteristics of a scheme is less clear, in particular because digital identity services may be offered free of charge (that is, on a non-commercial basis) to individuals, for reasons such as inclusion or perhaps even as a means of surveillance.

In these examples, the "✓" symbol indicates a characteristic that we would regard as positive from a privacy perspective, whilst the "×" symbol is used to indicate a negative characteristic.

### 5.2.1 Austria

| | |
|---|---|
| **Regulatory** | ✓Comprehensive[105] data protection law[106]<br>✓Independent data protection authority[107]<br>✓Personal data = any information relating to an identified or identifiable data subject. Includes facts that can be associated with a person, such as a person's name, date of birth, address, salary size, as well as value judgments.<br>✓Sensitive data = any information relating to a person's racial or ethnic origin, political opinions, religious or philosophical beliefs as well as health or sex life. |

[105] Applying to both the public and private sectors.
[106] Data Protection Act 2000
[107] Österreichische Datenschutzbehörde

| | |
|---|---|
| **Technical** | ✓ Data collected, held, processed and shared by the eID system: CCR number, name and date of birth<br>✓ Integration with other services: integrates with 12 services, including taxes, benefits, car registration, personal documents<br>✓ Numerous privacy features, including separation of identities by sector.<br>✓ Available in both card and mobile formats |
| **Commercial** | ✓ Government-led model, cards free to citizens, with few active threats to privacy, though vulnerable to passive threats |

### 5.2.2 Canada

| | |
|---|---|
| **Regulatory** | × No comprehensive data protection law, but rather a privacy law pertaining to the federal government,[108] a federal privacy law pertaining to the private sector,[109] and province-specific and sector-specific privacy laws.<br>✓Independent data protection authority[110]<br>✓Generally speaking, personal data = information about an identifiable individual, but ≠ contact information, including email addresses, which an organization collects, uses or discloses solely for the purpose of communicating with a person in relation to their employment, business or profession.[111]<br>✓Sensitive data is not defined under federal laws. However, the following data is treated as sensitive: medical records, income records and information about sexual orientation. |
| **Technical** | ✓Privacy a strong requirement in the hub design<br>✓De-coupling of credential service providers and service providers preventing "do not track" and linkability issues.<br>×In theory hub can track usage and link identities, however it is stateless[112] so in normal operation this should not happen. |
| **Commercial** | ✓Service providers pay for authentications with some value flowing to credential service providers (Banks)<br>✓ Commercial model linked to the issuance of a public services card that will reduce KYC costs to banks in future. |

[108] Privacy Act 1983
[109] The Personal Information Protection and Electronic Documents Act 2000, most recently amended by the Digital Privacy Act 2015.
[110] Office of the Privacy Commissioner Canada
[111] This situation was recently clarified in amendments made to PIPEDA (the federal private sector privacy legislation) by the Digital Privacy Act 2015: https://www.priv.gc.ca/resource/fs-fi/02_05_d_63_s4_e.asp.
[112] Meaning that no transaction-specific data is required to persist in the system.

### 5.2.3 Chile

| | |
|---|---|
| **Regulatory** | ✓ Comprehensive data protection law but falling short of EU standards[113]<br>× No independent data protection authority[114]<br>✓Personal data = any information relating to an identified or identifiable natural person.<br>✓Sensitive data = a person's physical or moral characteristics and facts or circumstances of their private life, such as personal habits, racial background, political opinions, religious beliefs, physical and mental health and sex life. |
| **Technical** | ✓ Data collected, held, processed and shared by the eID system: RUN[115] number, digitised photograph, signature, fingerprint, name, address and telephone number<br>✓ Integration with other services: standard OpenID approach, numerous relying parties in place<br>× No particular privacy preserving features included such as sector-specific identifiers or restrictions on usage.<br>× Nascent scheme, some troubles with deployment |
| **Commercial** | ✓ Government-led model, citizens pay for cards |

### 5.2.4 Ecuador

| | |
|---|---|
| **Regulatory** | × No data protection law, although a draft Bill was introduced in parliament in January 2015.<br>× No independent data protection authority. |
| **Technical** | × Data collected, held, processed and shared by the eID smart card system: Identification number, fingerprint code, names of the owner, the date and place of birth, nationality, gender, marital status, place and date of issue, expiration date, photo, holder's signature, signature of the competent authority, blood type and whether or not an organ donor<br>× Integration with other services: standard OpenID approach, numerous relying parties in place<br>× Technically ambitious, but structures unclear<br>✓ Nascent scheme |
| **Commercial** | ✓ Government-led model, cards free to citizens, with few active threats to privacy, though vulnerable to passive threats |

[113] Law on the Protection of the Private Life (Law No. 19,628) 1999
[114] Although a new law is being contemplated which will establish a data protection authority: http://www.privacylaws.com/Int_enews_21_1_15
[115] Rol Único Nacional – the national identification number.

### 5.2.5 Estonia

| | |
|---|---|
| **Regulatory** | ✓Comprehensive data protection law[116]<br>✓Independent data protection authority[117]<br>✓Personal data = any data concerning an identified or identifiable natural person. Personal data for the purposes of Estonian law includes both digital and paper based records.<br>✓Sensitive data = data on a person's political opinions, religious or philosophical beliefs, ethnic or racial origin, health records, genetic information, biometric data, sex life, trade union membership, criminal record or having fallen victim to an offence before relevant court procedures have taken place. |
| **Technical** | ✓Standard smart card technology<br>✓Mobile ID that supports online use only with limited data |
| **Commercial** | ✓Card readers widely distributed. |

### 5.2.6 Kenya

| | |
|---|---|
| **Regulatory** | × No comprehensive data protection law, but some regulation in the information and communication sectors.[118]<br>× No independent data protection authority. |
| **Technical** | ✓ Data collected, held, processed and shared by the eID system: name, gender, tribe or race, date of birth, place of birth, occupation, place of residence and postal address, finger and thumb impressions, date of registration.<br>× Integration with other services: information not yet available<br>× No outstanding privacy preserving features<br>× Nascent scheme, not yet begun issuing cards.<br>× Data from Kenya's National Registration Bureau shows that of the 915,101 applications for eIDs made in 2014, more than 175,000 were rejected.[119] |
| **Commercial** | ✓ Government-led model, cards free to citizens, with few active threats to privacy, though vulnerable to passive threats |

[116] Data Protection Act 2007
[117] Estonian Data Protection Inspectorate
[118] Kenya Information and Communication Act as read with the Kenyan Information and Communication (Consumer Protection) Regulations.
[119] Verah Okeyo, "The politics of identity cards in Kenya, and how registration law promotes sour ethnic divisions," *The Daily Nation,* 10 November 2015, available at http://www.nation.co.ke/lifestyle/DN2/Inside-Kenya-identity-crisis/-/957860/2949292/-/hg8eotz/-/index.html.

### 5.2.7 Malaysia

| | |
|---|---|
| **Regulatory** | ✓Comprehensive data protection law[120]<br>× No independent data protection authority<br>✓Personal data = information in respect of commercial transactions only that relates directly or indirectly to the data subject, who is identified or identifiable from that information or from that and other information in the possession of a data user.<br>× Sensitive data protections are restricted only to data related to health, political opinion, religious belief or criminal record. |
| **Technical** | × Data collected, held, processed and shared by the eID system: Name, address, race, citizenship, religion (for Muslims), fingerprints<br>× Integration with other services: national identity card, driving licence, passport information, health information, electronic purse, ATM access, transit application and the Public Key Infrastructure feature for online transactions.[121] But insufficient separation between services to ensure privacy.<br>× Uncontrolled access, personal information stored on the card |
| **Commercial** | ✓ Government-led model, citizens pay for cards |

### 5.2.8 Nigeria

| | |
|---|---|
| **Regulatory** | × No comprehensive data protection law; two separate Bills have been pending since 2008 and 2010 respectively. However, the National Information Technology Development Agency issued in 2013 Draft Guidelines on Data Protection "the Guidelines".[122]<br>× No independent data protection authority<br>✓Under the draft Guidelines, personal data = any information relating to an identified or identifiable natural person; information relating to an individual, whether it relates to his or her private, professional or public life, including "anything from a name, address, a photo, an email address, bank details, posts on social networking websites, medical information, or a computer's IP address"<br>✓Under the draft Guidelines, sensitive data = data relating to religious or other beliefs, sexual orientation, health, race, ethnicity, political views, trade union membership, criminal record. |

[120] Personal Data Protection Act 2010
[121] Yap Ai Kee, Yeoh Choo Nee, Leau Yu Beng, and Tan Soo Fun, "Security Issues on Identity Card in Malaysia," IACSIT International Journal of Engineering and Technology, Vol. 4, No. 5, October 2012, available at http://www.ijetch.org/papers/445-R30008.pdf.
[122] http://www.nitda.gov.ng/documents/Guidelines%20on%20Data%20Protection%20Final%20Draft3.5%20Final.pdf

| | |
|---|---|
| **Technical** | × Data collected, held, processed and shared by the eID system: personal details, address, parents, next of kin, origin, identification documents, disability data, signature and biometrics<br>× Integration with other services: transit, health, pension, banking. But insufficient separation between services to ensure privacy.<br>× MasterCard branding, complaints about citizens not receiving or collecting their cards.<br>× The Nigerian government recently warned against approaching unauthorised agencies and individuals for national identification registration, after reports of unauthorised fake eID cards being registered.[123] |
| **Commercial** | ✓ Government-led model, cards free to citizens, with few active threats to privacy, though vulnerable to passive threats |

### 5.2.9 Pakistan

| | |
|---|---|
| **Regulatory** | × No data protection law, although a draft Electronic Data Protection Act was introduced in parliament in 2005.<br>× No independent data protection authority.<br>✓ Under the draft Act, personal data = any information relating to an individual, identified or identifiable, directly or indirectly by reference to any other information.<br>✓ Under the draft Act, sensitive data = data revealing racial or ethnic origin, religious, philosophical or other beliefs, political opinions, membership of political parties, trade unions, organisations and associations with a religious, philosophical, political or trade-union, or information as to the health or sex life of an individual and financial, or proprietary confidential corporate data. |
| **Technical** | × Data collected, held, processed and shared by the eID system: name, gender, father's name (husband's name for married females), identification mark, date of birth, national identity card number, family tree ID number, address, permanent address, religion, signature, photo, fingerprint.<br>× Integration with other services: transit, immigration, police. But insufficient separation between services to ensure privacy.<br>× Numerous privacy concerns.<br>✓Used to facilitate cash benefit to IDPs and flood victims |
| **Commercial** | ✓ Government-led model, citizens pay for cards |

[123] "Military alerts Nigerians on circulation of fake national identity card," *Premium Times,* 23 September 2015, available at http://www.premiumtimesng.com/news/top-news/190532-military-alerts-nigerians-on-circulation-of-fake-national-identity-card.html.

### 5.2.10 Peru

| | |
|---|---|
| **Regulatory** | ✓Comprehensive data protection law[124]<br>✓Independent data protection authority[125]<br>✓Independent identification authority (RENIEC)<br>✓Personal data = any information on an individual which identifies or makes him identifiable through means that may be reasonably used.<br>✓Sensitive data = biometric data, data concerning racial and ethnic origin; political, religious, philosophical or moral opinions or convictions, personal habits, union membership and information related to health or sexual life. |
| **Technical** | × Data collected, held, processed and shared by the eID system: name, unique ID number, date of birth, marital status, photograph, fingerprint, voting number.<br>× Integration with other services: unknown.<br>× Currently digital smart card is small scale, concerns around the price, could exclude poor.<br>✓ Won award for card security features and functions. |
| **Commercial** | ✓ Government-led model, citizens pay for cards |

### 5.2.11 Saudi Arabia

| | |
|---|---|
| **Regulatory** | × No data protection law, some sectoral regulation.<br>× No independent data protection authority |
| **Technical** | ✓Based on standard smart card and digital identity technology<br>×Believed to be a highly centralised service with significant data pools<br>×Stores more data that many schemes including data about family and religion. |
| **Commercial** | ✓Government-led model, cards free to citizens, with few active threats to privacy, though vulnerable to passive threats |

[124] Data Protection law (Ley N° 29733)
[125] National Authority for Personal Data Protection

### 5.2.12 United Kingdom

| | |
|---|---|
| **Regulatory** | ✓Comprehensive data protection law[126]<br>✓Independent data protection authority[127]<br>✓Personal data = any data from which a living individual may be identified, either alone from that data or in conjunction with other information already in the possession of, or which is likely to come into the possession of, the person who determines the purposes for which personal data will be processed provided that, under current case law, such data is biographical of or focuses on the Data Subject.<br>✓Personal data ≠ IP addresses[128]<br>✓Sensitive data = any information relating to a person's racial or ethnic origin, political opinions, religious or philosophical beliefs, membership of a union, physical or mental health, sexual life or information about the commission of offences. |
| **Technical** | ✓Significant effort put into end-user experience.<br>✓Decoupling of IDPs and service providers.<br>×In theory hub can track usage and link identities, however it is stateless so in normal operation this should not happen. According to research,[129] in both the Verify and FCCX systems, "the hub – and whoever can gain control over it, legitimately or not – can link transactions of the same user, as defined by a user account at an IDP, across different [service providers]… giv[ing] the hub excessive visibility into the activities of all users", so that if actively compromised the hub could be used to impersonate users and gain access to their accounts.<br>×Inclusion of "matching data set" in all identity transactions increases potential for tracking and surveillance. |
| **Commercial** | ✓Aim is to create market place for digital identity<br>✓Government is seeding the service by being the first user.<br>×If system does not grow fast enough or is not sufficiently relevant to private sector, it will not be commercially viable for IDPs. |

[126] Data Protection Act 1998
[127] Information Commissioner's Office
[128] Information Commissioner's Office, Personal information online code of practice, https://ico.org.uk/media/for-organisations/documents/1591/personal_information_online_cop.pdf
[129] Luis T. A. N. Brandao, Nicholas Christin, George Danezis, and Anonymous, "Toward Mending Two Nations-Scale Brokered Identification Systems" (2015) 2 *Proceedings on Privacy Enhancing Technologies,* 1-22, available at http://www0.cs.ucl.ac.uk/staff/G.Danezis/papers/popets15-brokid.pdf

# 6 CURRENT DIGITAL IDENTITY ARCHITECTURAL MODELS

There are numerous approaches to digital identity in the marketplace. In order to provide a means of comparing the qualities and characteristics of these schemes, we have identified seven archetypal models of digital identification providers. The focus of this section is to provide an introduction to these models, in consideration of the privacy and other aspects of the various approaches. Sections 7 and 8 consider in detail the privacy risks of each model and opportunities for mitigation.

The models are presented in a sequence from more centralised to more decentralised. In general terms, the more centralised models prioritise user experience and inclusion over privacy. The more decentralised models prioritise privacy over user experience and inclusion. There are nuances with each model, which are explored in sections 5 and 7.

### 6.1.1 Monolithic internet identity provider

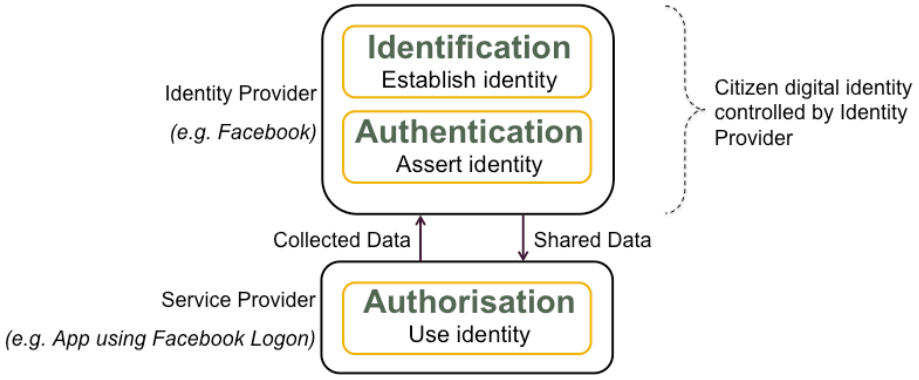



*Figure 11: Monolithic Internet Identity Provider*

**Examples: Facebook, Google**

The internet giants (Facebook and Google in particular) have built businesses around collecting information about consumers that can be used to profile customers for targeted advertising purposes. Initially information was collected from usage of the platforms, e.g. the posts made to the site or the information searched for, as well as self-asserted identity information. Logon services enable them to collect and share information with third parties, extending the presence of their brands as well as providing additional data for their profiling.

The digital identity schemes offered by the internet giants are monolithic. Whilst they are similar to open protocols such as OpenID Connect, which are designed for a more open federated approach to identity, the internet giants act as single universal identity schemes. The user is given the choice of logging on with Facebook, as opposed to being given the choice of performing a federated logon where Facebook is one of many federated identity providers.

The digital identity offered by these services is generally considered low assurance. Identification is initially self-asserted – based on information the user enters into their user profile. The identity providers often insist on a 'real names' policy, despite known problems[130]

[130] http://geekfeminism.wikia.com/wiki/Who_is_harmed_by_a_%22Real_Names%22_policy%3F

and publicity backlashes[131]. Despite claims that the policies are intended to increase customer trust as they will "know" who they are dealing with, a key driver is likely to be attempts to monetize named individuals. This data can be complemented and corroborated through data received from service providers (e.g. online merchants), usage and social graph data, as well as verified accounts for brands and celebrities[132]. This data is increasingly being used to support a higher level of assurance in identity. Start-ups such as Trulioo and Veridu actively use social media to support identification of individuals. OIX in conjunction with the UK government's Verify programme have published a paper on the use of such data to complement traditional identity verification methods.[133]

There is no separation of identification and authentication in these services. A user asserts their digital identity by performing a logon to the internet giant's site. The assertion of identity is inextricably linked to the account that the user has with the internet giant.

These logon services may include prompting the user for consent to share data. The granularity and control around such consent is increasing. Facebook provides guidance to app developers on the use of permissions[134] (the means through which consent is gained). This includes only asking for permissions that are needed. Similarly the latest version of the Android operating system ("Marshmallow") allows permissions to be requested "just in time". Previously an app would have needed to request all permissions up front. The guidance provided by Facebook is not mandated, and therefore presumably not enforced. App providers may still request consent to share a greater amount of data than is strictly necessary. They may also adopt an "all or nothing" approach where if you do not provide consent they do not allow access to the app.

### 6.1.2 Federated internet identity providers

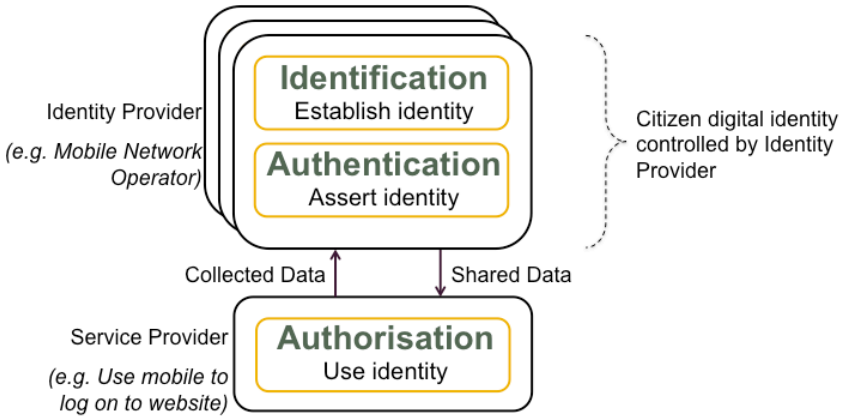



*Figure 12: Federated internet identity providers*

**Examples: GSMA Mobile Connect, Open ID Connect, Mobile Operators, PayPal, Amazon[135]**

[131] http://www.theguardian.com/technology/2015/nov/02/facebook-real-name-policy-protest
[132] https://en-gb.facebook.com/help/196050490547892
[133] http://openidentityexchange.org/2014/02/internet-life-verification-using-social-data-as-part-of-the-identity-verification-process/
[134] https://developers.facebook.com/docs/facebook-login/overview
[135] PayPal and Amazon are in this list because they are less dependent (than Facebook and Google) on data aggregation for their revenues, giving them more scope to support better privacy.

The OpenID Connect[136] specifications define open protocols for federated digital identity. The intention is to provide the end user with the ability to use an identity provider of their choice to logon to and share data with service providers. Technically the protocols are very similar to those used by the internet giants. In fact both Facebook Connect and OpenID Connect build on top of OAuth – a generic protocol for granting authorisation to access resources on the internet.

The specifications include support for the user to provide consent for the data that is shared along with the logon. They include support for aggregated and distributed "claims" (i.e. attributes).

OpenID Connect has been implemented by many organisations including Google, Microsoft and Salesforce.com[137]. More significantly for this study the GSMA's Mobile Connect[138] programme uses the OpenID Connect specifications. Mobile Connect is a major undertaking by the GSMA and its members with the objective of establishing secure digital identities and consented data sharing, leveraging the security afforded by the SIM. The GSMA are focusing their efforts on a number of markets including India, with the obvious potential linkage to Aadhaar.[139]

### 6.1.3 State issued eID cards

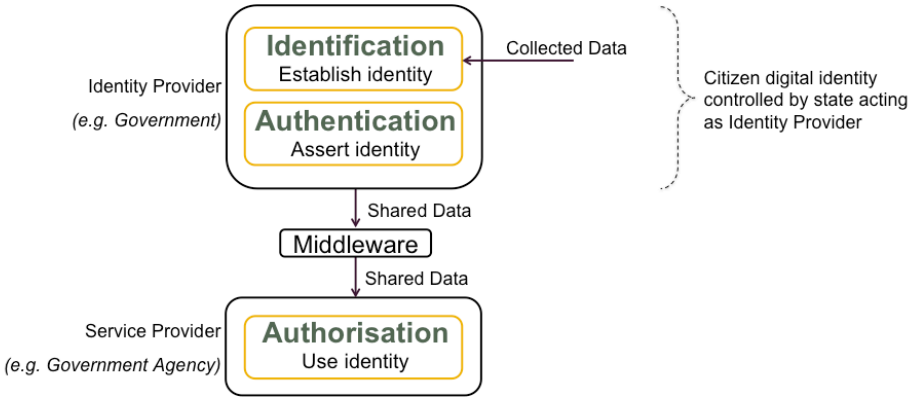



*Figure 13: State issued eID cards*

**Examples: Estonia eID, Nigeria eID and numerous other national eID schemes – also RealMe, the online-only digital identity issued by the New Zealand Government.**

Numerous states have issued eID cards[140]. In many cases these represent an evolution of paper-based identity cards. Citizens are typically issued with a smart card that acts as an authentication token as well as potentially holding some attribute data. In some markets (e.g. Estonia) mobile versions of the eID exist where the SIM (which is itself a smart card) is used in place of the card form factor.

State issued eID systems are generally proprietary, to the extent that they are created by governments addressing the particular nuances of the country concerned. So whilst the underlying technology is often common (e.g. ISO/IEC 7816 electronic identification cards) there will often be differences both in terms of the data being managed and the systems providing that management. Furthermore the governance of such systems will vary as well. In some

[136] http://openid.net/connect/
[137] http://openid.net/2014/02/26/the-openid-foundation-launches-the-openid-connect-standard/
[138] https://developer.mobileconnect.io/docs-mobile-connect-api
[139] http://telecom.economictimes.indiatimes.com/tele-talk/mobile-as-the-digital-identity-for-india/954
[140] According to the World Bank ID4D dataset 113 states currently have eID systems

cases (e.g. Peru) an autonomous entity runs the scheme, in other cases (e.g. Saudi Arabia) a specific government department is responsible.

Identification for eID schemes will be linked to registration of the citizen with the government, whether at birth, at another significant juncture (such as entering adulthood) or through a new process where no suitable register already exists (as was the case in India).

Authentication is performed using the smart card (or SIM) personalised with citizen data and the necessary digital certificates and cryptographic keys.

Usually both identification and authentication are performed as part of the scheme and hence the state can be viewed as the "identity provider".

The usage of state issued eID varies as well, ranging from limited government use to broader support for commercial use. The majority of schemes support both logon (or "authentication") and digital signing. Integration of smartcard eID for online services often involves the use of a smart card reader and middleware running on the citizen's PC, to provide applications with a means of accessing the card. Other schemes also include an infrastructural middleware layer connecting the core eID services to service providers (as is the case in Estonia and represents the model promoted in the World Bank's ID4D research).

The data processed by state-led eID schemes varies too. Some provide core attributes only (name, date of birth etc). Others provide data and application containers for a range of public services including health and transit.

**BankID (Norway)**

We consider that BankID in Norway fits under this model. Whilst BankID is issued by the banks, in many respects it is equivalent to a national eID in terms of both the identity technology and usage. BankID is a PKI[141] issuing digital certificates to citizens that can be used (in conjunction with the associated cryptographic keys) to perform digital signatures and include basic attributes about the citizen. There are two main differences from the majority of eID schemes:

- The digital certificates are issued by the citizen's bank. The banks can however be viewed as outsourced suppliers to the government.
- The citizens' cryptographic keys are stored in the cloud, with access to those keys controlled through OTP based authentication.

In other respects the system will be very similar to a national eID scheme. Digital certificates are issued by a certificate authority and certificate validity is determined by checking both the certificate expiry (contained within the certificate) and checking an OCSP (Online Certificate Status Protocol) server.

[141] Public Key Infrastructure

### 6.1.4 Brokered Identity Providers (IDPs)

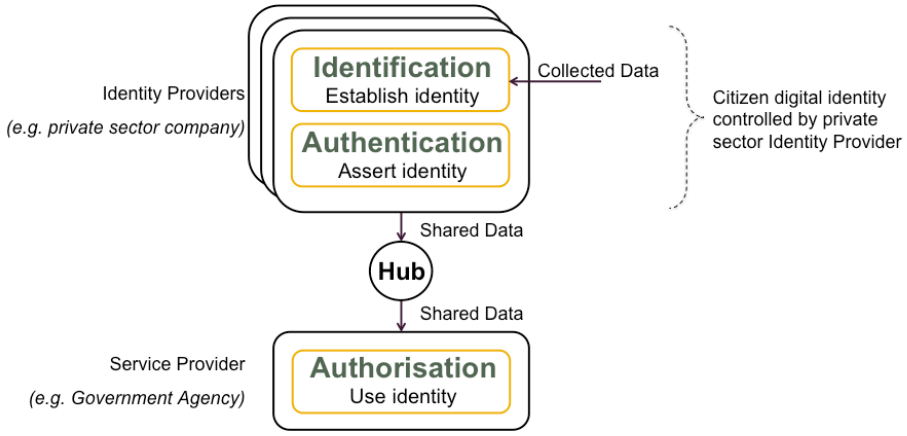



*Figure 14: Brokered IDP*

**Examples: UK Verify, US Connect.Gov**

The UK government's Verify programme is seeking to establish a private sector marketplace for digital identity. The ultimate goal is that private sector organisations will create and manage digital identities on behalf of citizens and that citizens will be able to use those identities for a wide variety of services, including government. In this way the government will share the cost of digital identity with the private sector.

At the time of writing, the Verify programme consists of a number of private sector identity providers who are connected via a hub to a growing number of government digital services. Ultimately the intention is that non-government online services will also be able to leverage the same identities – potentially via a private sector hub.

The UK government are hoping to establish this as a model to be replicated in other markets, bringing further economies of scale through standardisation. In particular, core assets from the UK scheme have been offered "open source" to Australia's recently formed Digital Transformation Office (DTO) which is in the process of establishing a similar initiative in Australia.

Although not directly relevant to the description of the model, Verify has had a significant focus on user experience and on inclusion with projects being run to explore the issues of providing digital services (of which digital identity is the key enabler) to individuals who are at present financially or digitally excluded. Whilst many of these issues face individuals who are UK residents, this is not exclusively the case. Citizens of other countries, for example, may be entitled to UK pension payments[142] through family connections.

Verify is focused on "logon" as opposed to digital document signing. In that sense it is similar to the federated logon schemes such as OpenID Connect and Mobile Connect. Federated logon schemes however do not necessarily require a hub in order to function. OpenID Connect has optional support for an identity provider discovery mechanism which, in theory, should allow service providers to redirect the user to their identity provider without the need for a centralised brokerage function. We therefore use the term "brokerage" as that more accurately describes the architecture of the Verify service.

[142] http://oixuk.org/?p=1793

In this model identification and authentication are performed by the organisation, the "identity provider". There will potentially be many identity providers that the citizen can choose from.[143]

The identification process involves obtaining and verifying identity evidence from a variety of sources including authoritative sources, breeder documents, credit reference agencies, supporting documents, as well as checking the activity history associated with the identity in order to detect anomalous behaviour.

When an identity is asserted, the data to be shared is delivered via a hub which performs the function of standardising the messaging as well as routing to the relevant service provider.

In the Verify programme identity is defined as the combination of name, date of birth, postcode and gender. This data is passed in all interactions, as a central tenet of the Verify design is that service providers should be able to automatically "match" identity requests with accounts in their systems without the need for separate identification or binding steps.

The hub performs a level of decoupling between service providers and identity providers. Identity providers are unable to see which service provider their identity assertion is being delivered to. Service providers are unable to see which identity provider made the assertion. They are simply provided with an assertion which states that an approved identity provider, who has been certified to the required standard, has provided the assertion.

### 6.1.5 Brokered Credential Service Providers (CSPs)

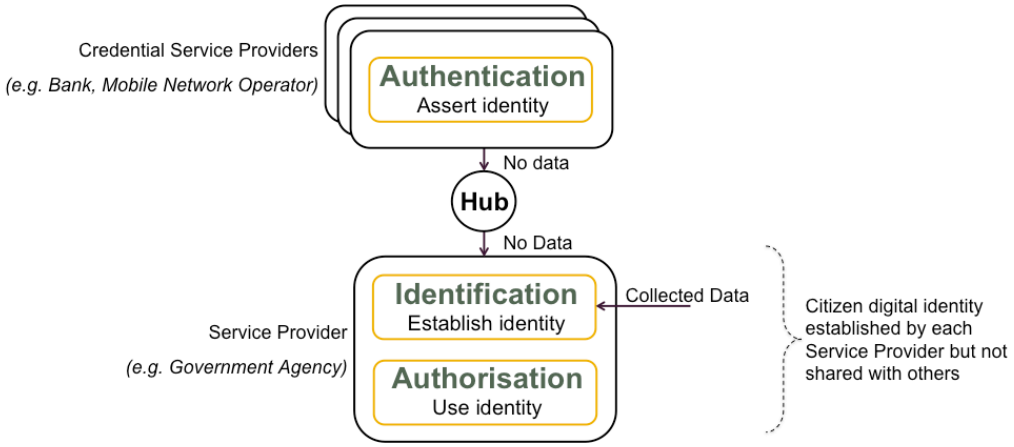



*Figure 15: Brokered CSPs*

**Examples: Canada Credential Broker Service (believed to be unique)**

The Canadian credential brokerage service (operated by SecureKey) takes an even more decentralised approach to identity. In fact the scheme does not provide identification services at all. Rather it is down to the individual service provider to perform the necessary ID proofing or "Know Your Customer" steps in order to link the authentication credential with the relevant identity (or citizen) within the service provider's systems.

The FIDO Alliance has a similar approach, providing a mechanism for an end-user to establish authentication credentials on their mobile device. The FIDO specifications do not define an explicit hub per se rather FIDO client software resident on the end-user's device provides

[143] Verify explicitly allows citizens to set up a number of identities, up to one per identity provider.

integration between the authentication credential and the service provider. The end result is the same, namely that the end-user has an authentication credential that allows them to repeatedly assert their "identity" although the linking of that identity to actual attribute data is down to each service provider. No attribute data is passed in the authentication process.

In both of these examples, the main benefit is the removal of issues associated with the fragmented approach to password management that exists today. Instead of the citizen needing to manage multiple passwords themselves, they are able to use a single credential to logon to multiple services.

Both FIDO and the Canadian scheme include a strong element of unlinkability. In both schemes, a service provider receives a service provider-specific anonymous identifier that is linked to the authentication token allowing the service provider to establish that an authentication is from the same identity as a previous transaction. If multiple service providers were to compare authentication data they would not be able to link identities together, at least not using the authentication data that is passed.

Clearly such an approach does not solve the more fundamental issue of establishing digital identity in the first place. It does however provide a strongly private way to assert, in digital contexts, an identity already established.

It would also appear to be consistent with some of the ideas laid out in the World Bank's ID4D initiative, which advocates a separation of authentication credential and unique identifiers from attribute data.

### 6.1.6 Personal IDP

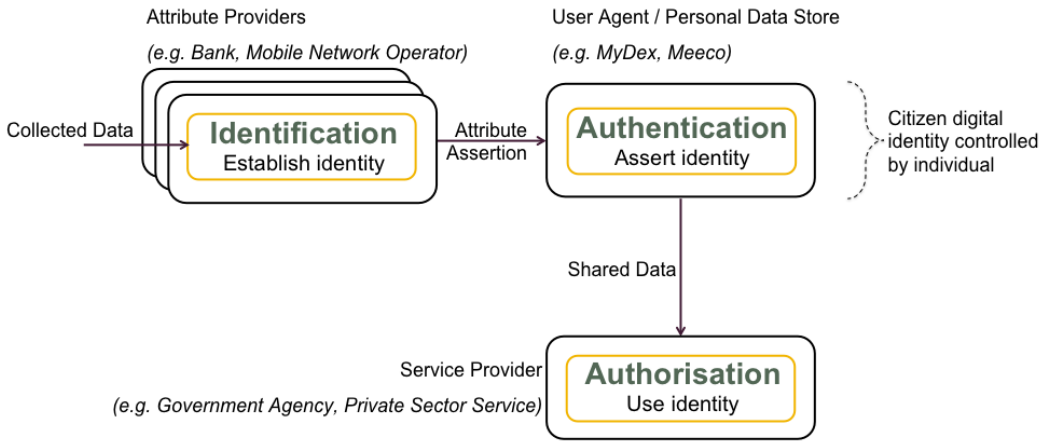



*Figure 16: Personal IDP*

**Examples[144]: MyDex, Meeco, Microsoft u-Prove**

[144] See http://pde.cc/directory/ for an up to date list.

In response to privacy concerns regarding some approaches to identity and as a reaction to the way businesses acquire and profile consumer data, there is a section of the identity industry that focuses on person-centric identity. There are two primary approaches to making digital identity person-centric:

- Through the use of personal data stores, individuals are given much greater control over their data than would otherwise be the case. Initiatives such as the Internet Identity Workshop (led by "Identity Woman"[145], Kaliya Hamlin) and Project VRM[146] (led by Doc Searls) promote an approach where citizens manage their service providers, not the other way around.
- Privacy-preserving claims-based identity technology being developed by Microsoft[147] and IBM[148] and the focus of a recent EU funded project ABC4Trust[149].

These approaches make the citizen the gatekeeper for their attribute data, bringing high levels of control to the individual. As yet none of these initiatives has reached significant commercial scale, however they provide a useful comparison to other approaches that are less user-centric.

Mydex (a personal data store provider), in particular, was one of the original identity providers to the UK government's Verify programme, although they have recently taken on a new more consultative role in that programme.[150]

### 6.1.7 No IDP

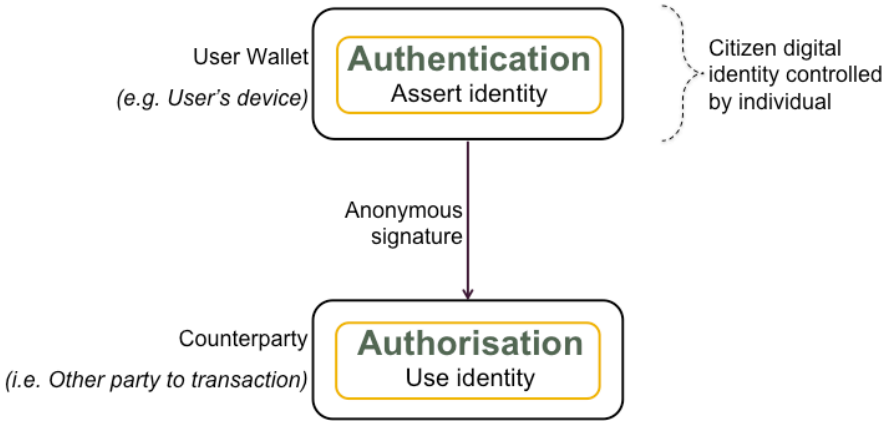



*Figure 17: No IDP*

**Examples: Bitcoin, various Blockchain identity startups such as Shocard**

Bitcoin is a decentralised and unregulated platform for exchange of the Bitcoin currency. The Bitcoin system itself and the emerging blockchain technologies promote highly decentralised and anonymous forms of identity[151]. In Bitcoin itself, the user's identity (their "address") is self-

[145] http://www.identitywoman.net/
[146] http://cyber.law.harvard.edu/projectvrm/Main_Page
[147] http://research.microsoft.com/en-us/projects/u-prove
[148] http://www.zurich.ibm.com/idemix/
[149] https://abc4trust.eu/
[150] https://identityassurance.blog.gov.uk/2015/03/25/gov-uk-verify-and-mydex/
[151] https://www.gov.uk/government/publications/distributed-ledger-technology-blackett-review

generated and self-asserted, with responsibility for the creation and protection of the cryptographic keys used to assert the identity lying entirely with the user.[152]

The blockchain community is in the process of exploring many issues around identity, such as how KYC data can be irrevocably linked to a financial transaction conducted on an openly readable ledger but remain private. Through cryptographic techniques such as Zero Knowledge Proof assertions, such an entitlement to enter into a transaction may be possible without revealing the actual attribute data.

Whilst these developments are extremely nascent, Blockchain technology is currently receiving a great deal of attention (and investment), especially from the banking community.

[152] In reality most individuals use wallets provided by third parties.

# 7 RISKS

The recognised way of determining privacy risks in any service is a privacy impact assessment[153]. This involves conducting a risk assessment focusing on the privacy aspects of the service. The actual privacy risks present in a service will arise from the specifics of the system or service. The scope of the service will determine the threats – who may attempt to undermine the service. The technology employed and how it is used will determine the vulnerabilities – what weaknesses in the system may give the threat agent the opportunity to compromise it. Together the threats and vulnerabilities constitute risks.

A service may contain vulnerabilities but if there are only limited threats (e.g. the current scope of the service is too small to be of interest) the overall risk may be low. Conversely a service may make robust use of technology but due to its scale and scope be a target of great interest, resulting in the risk being higher. Risk arises from a combination of incentive and means. This includes risks arising from passive threats, such as incompetence and negligence.

When considering digital identity schemes, obtaining a compelling assessment of risk requires analysis of the specifics of each scheme. Different schemes may use similar technology in different ways resulting in markedly different risk profiles. For example, smart cards are generally viewed as being a secure technology for holding end-user authentication credentials. They typically contain tamper resistant hardware for the storage of cryptographic keys and the secure execution of cryptographic operations. The security of such a component will however be dependent on many factors including the length of cryptographic keys and the algorithms employed, the frequency with which keys are replaced, the PIN management used to protect access to the keys and the protections the card provides against attempts to infer secret data.[154]

In order to provide a meaningful assessment of the privacy risks associated with different digital identity architectures, this document includes a general assessment of risk, considering the types of threat and vulnerability that may apply. As far as is possible within the scope of this document, a view of which architectures are likely to perform better from a privacy risk perspective is also provided. This is no substitute for performing a privacy impact assessment specific to a particular digital scheme, but the following subsections provide a means for comparing at a high level the likely privacy characteristics of different approaches.

## 7.1 Threats

As for any other digital service, the threats to a digital identity scheme fall into one of three broad classes:

- Threats to the confidentiality of the scheme. These are direct threats to the privacy of the enrolled participants, and to confidential data relating to the service itself;
- Threats to the integrity of the scheme – that is, where the data held by the scheme is altered by an unauthorised party;

[153] See http://www.oaic.gov.au/images/documents/privacy/privacy-resources/privacy-guides/guide-to-undertaking-pias.pdf

[154] E.g. through differential power analysis

- Threats to the availability of the scheme. These are threats to the service itself, disrupting either enrolment or the subsequent provision of identity-related services, such as identity assurance.

These threats might further arise from a number of sources, such as:

- Malevolent threats (entities attacking a digital identity service, to cause disruption by undermining its integrity, or undermining the privacy of individuals);
- Legitimate businesses that potentially compromise the privacy of personal information through their legal business practices;
- Passive threats to the service, which arise through either incompetence or negligence on the part of those responsible for its provision.

The range of threats a digital identity service might be exposed to is summarised in the following table. For each threat, a potential severity is given, as follows:

- **High**: Threat of direct disruption or financial loss to citizens
- **Medium**: Threat of embarrassment, inconvenience or annoyance to citizens
- **Low**: Minimal threat.[155]

| Threat Type | # | Example | Severity |
|---|---|---|---|
| **Confidentiality threats** | | | |
| Malevolent | T.1 | Identity theft: systematic exposure of an individual's personal information allowing subsequent impersonation of the identity. | High |
| | T.2 | Data breach: exposure of a large amount of personal information pertaining to a large number of citizens by a criminal gang enabling wide-scale fraud or extortion. | High |
| | T.3 | Individual surveillance: exposure of personal information (including usage data) that would reasonably be assumed to be private, to stalker, private investigator etc. | Medium |
| | T.4 | Mass surveillance: exposure of large amount of personal information pertaining to a large number of citizens (including usage data) that would reasonably be assumed to be private. | High |
| Inappropriate | T.5 | Passing of personal information (with citizen's consent) to unvetted parties: exposure of personal information if third party does not behave appropriately. | Medium |
| | T.6 | Repurposing of data: e.g. data collected for health reasons shared with tax authority. | Medium |
| | T.7 | Personal information made public: publication of large volumes of data (with citizen's consent) that can be aggregated and mined resulting in the privacy of individuals being undermined. | Medium |

[155] In a full risk assessment, threats would be assessed in terms of the potential "impact" to the service and the potential "gain" to the attacker. For the purposes of this document a severity rating is provided which is a high level indication of the seriousness of the threat.

| Threat Type | # | Example | Severity |
|---|---|---|---|
| Passive | T.8 | Poor operational security: leading to private data being exposed. Although this is a possibility for any scheme, the lack of appropriately secured data centres in many of the emerging economies makes this more likely. For example, it is commonplace in many countries to see data centre security doors propped open, in order to ease access for operators, without thought for the security implications. | High |
| | T.9 | Poor operational processes: for example, turning on data logging in order to investigate a problem, and then neglecting to turn it off again once the problem has been isolated. | High |
| | T.10 | Changing security landscape: failure to review security protocols, key lengths etc at regular intervals, so that a malevolent attacker is more likely to succeed. | High |
| **Integrity threats** | | | |
| Malevolent | T.11 | Attacks on the cryptographic underpinnings of a scheme; access to scheme private keys, for example, so that authentication services can no longer be trusted | High |
| | T.12 | Alteration or deletion of data: this could be to disrupt the service, bring it into disrepute, or to enable identity takeover. | High |
| Inappropriate | T.13 | None identified | Low |
| Passive | T.14 | None identified | Low |
| **Availability threats** | | | |
| Malevolent | T.15 | Denial of service attacks, so that authentication services are not available | Medium |
| | T.16 | Attacks on the cryptographic underpinnings of a scheme; deletion of scheme private keys, for example, so that authentication services can no longer be provided, and new registrations cannot be supported | High |
| Inappropriate | T.17 | None identified | Low |
| Passive | T.18 | Poor attention to the need to renew cryptographic certificates and other elements of the service such as software licences, so that the service becomes untrusted by automated services or simply unavailable. | Medium |

*Table 6: Categories of Threats*

It is not unknown for these threats to be actualised in combined attacks, so that for example a denial of service attack on a digital service might be intended to distract a service provider sufficiently so that a more subtle attack on the confidentiality or integrity of a service isn't noticed until after the fact.

## 7.2 Vulnerabilities

Vulnerabilities are weaknesses that could lead to the compromise of confidentiality, availability or integrity of the information assets being managed by a system. For the purposes of this document we provide a high level view of the types of vulnerability that may exist in digital identity systems, particularly focusing on vulnerabilities that result in the privacy of citizen data being reduced.

### 7.2.1 Vulnerability types

The following table lists types of vulnerability that may exist in a digital identity system. We have grouped these under the DIPPs established in Section 9.2. For each type of vulnerability, we provide examples of different severities of vulnerability (high, medium and low severity). In the next section we then use this measure to score the potential vulnerability of the digital identity architectural models.

| Vulnerability type | High | Medium | Low |
|---|---|---|---|
| **V.1 Fair and lawful processing** | | | |
| Data held in jurisdictions or organisations with different standards | Digital identity data duplicated across multiple organisations. | Digital identity data normalised[156] | Digital identity held by one organisation or in one jurisdiction with established standards.[157] |
| Robustness of certification – self certification, privacy focused or not. | Self certified with no explicit PIA required. | Self certified but includes a PIA | Certification includes PIA from qualified third party |
| **V.2 Adequacy and quality** | | | |
| Poor data quality | Self asserted data only | Data checked against a small number of external sources | Data checked against authoritative sources |
| Static long lived identities | Identity re-issued / re-verified infrequently | Identity re-issued infrequently but established revocation lists checked. | Identity re-verified frequently using multiple reference points |
| **V.3 Explicit and legitimate purposes** | | | |
| Unnecessary data collection | Data collected without having a clearly defined purpose for the provision of digital identity | Approach to digital identity requires collection of significant amounts of data (e.g. transactional data) | Approach to digital identity minimises type and amount of data collected |
| **V.4 Minimal disclosure for a constrained use** | | | |
| Linkable identifiers | Single identifier used for many purposes | Sector specific identifiers | Organisation specific identifiers[158] |
| Lack of blinding | Other parties in transaction visible | Parties in transaction separated by hub | Information about other parties in transaction blinded cryptographically |

[156] i.e. data duplication minimised, ideally with each attribute managed by one organisation (although could be multiple organisations).
[157] Note this is in conflict with desire to minimize honeypots. Ultimately it is likely to be better to avoid honeypots although this makes the task of ensuring consistent application of standards more difficult.
[158] A further level of granularity (transaction specific identifiers) may also be possible.

| Vulnerability type | High | Medium | Low |
|---|---|---|---|
| Lack of pseudonymisation | Large amounts of data potentially shared in single assertions | Minimum data set required for all transactions but other data optional | Granular partial identity supported |
| Lack of "verification mode" services[159] | Source data attribute shared | Data is queried rather than shared | Data is queried and response bound to current transaction so cannot be reused |
| Biometric data shared or stored in central database | Biometric template shared or stored centrally | Only hash of biometric shared or stored | Biometric data never leaves reader device. Only authentication result shared. |
| **V.5 Openness and transparency** | | | |
| Lack of transparency | Unclear what data is collected | Data to be collected is defined but have to trust provider not to collect more | Data to be collected is defined and the system limits the need to trust the provider |
| Implicit data collection | Data collected that citizen may not be aware of | Data collected but citizen likely to be aware | Data collection requires deliberate citizen action |
| **V.6 Individual ownership and control of data** | | | |
| Weak authentication: Authentication token has potential for theft or loss or cloning | No notification of data collection | Has option for strong authentication but not mandated | Strong authentication including smart card equivalent or strong biometrics |
| Gap between user and credential (cloud hosted versus card) | Credential (e.g. signing key) is hosted in cloud with access to it controlled by password or software | Credential (e.g. signing key) is hosted in cloud with access to it controlled by hardware device | User credential is a physical device or biometric under their direct control |
| Unencrypted data readable without access controls Access to data possible by unauthorised / unauthenticated parties | Significant amounts of data visible | Some data readable (e.g. from user device/card) with no controls | Reading data (e.g. from user device/ card) only permitted by authorised parties |
| Non-expiring user credentials | Credentials not periodically replaced / updated | Credentials not systemically replaced / updated but confirmation from user periodically obtained | Credentials replaced at a frequency determined by risk assessment or security policy |

[159] i.e. instead of "Is person over 18?" ask "what is person's age"

| Vulnerability type | High | Medium | Low |
|---|---|---|---|
| **V.7 Accountability and auditing** | | | |
| Lack of audit | Unregulated or where no clear digital identity guidelines exist. | Self regulated with clear guidelines relevant to digital identity. | Service operated as part of regulated industry. |
| Access to data not audited | No auditable record of access to data | Auditable record of access to data held but not usually shared. May need court injunction. | Full audited record of access to data held centrally. |
| **V.8 Consent** | | | |
| Cached consent: Consent once, use many<br>Disclosure of personal information without explicit consent | Generic and forced opt-in for unspecified data collection and sharing | Single opt-in but for collection and sharing amongst clear and limited set of organisations | Citizen notified whenever data that is to be used for digital identity purposes is collected.[160] |
| No facility to withdraw consent, lack of revocation | No straightforward mechanisms to withdraw consent once it has been given | Mechanism provided to manage consent but cumbersome or complex to use | Intuitive and clear tools provided to manage consent. |
| **V.9 Data minimization and avoidance of honeypots** | | | |
| Pools of data | Large pool of data | No large pool of data in design but potential to collect data large pool of data | No ability to collect large pool of data |
| **V.10 Sensitive data** | | | |
| Inferred data (e.g. surname being used to infer religious or ethnic group) | Identifiers revealed can be used to infer multiple data items (e.g. name, date of birth encoded into identifier) | Identifiers revealed can be used to infer single items of data | Identifiers are random and impersonal |
| **V.11 Avoid exclusion** | | | |
| High cost to individual or lack of availability | Citizen required to pay, requires specific technology (e.g. Smart phone), requires mobile or bank account | Citizen not required to pay but requires specific technology (e.g. Smart phone) or requires mobile or bank account | Free at point of use and no specific hardware / technology requirement |

[160] Note that digital identity services may also, for example, collect usage data for monitoring performance. We are assuming in this category that strict rules apply to the collection and usage of such data, so that no unwanted leakage occurs.

| Vulnerability type | High | Medium | Low |
|---|---|---|---|
| **V.12 Restrictions on transfer and disclosure of data** | | | |
| Lack of restrictions | No restrictions in place or vetting of recipients, with onus being placed on citizen to provide consent | Data only shared with trusted third parties | Limited or no data sharing |
| **V.13 Pluralism and interoperability of systems and technologies** | | | |
| Proprietary technologies | System is proprietary and bespoke, leading to interoperability issues and also higher chance of hidden security issues | Some aspects of the system proprietary or provided by vendor seeking to establish lock-in | System built around open and widely used technologies |
| **V.14 Minimise the human element** | | | |
| Reliance on human action (Greater potential for incompetence and negligence to affect system) | Significant manual procedures around registration and other lifecycle events | Initial enrolment manual but once established other lifecycle events are automated. | System fully automated |
| **V.15 Robustness of technology** | | | |
| Complexity increasing likelihood of security weaknesses being present | Complex system with large scope | Simpler architecture but large scope | Simple system with a constrained scope. |
| Susceptibility to malware | Citizen identity relies on personal computing device software | Citizen identity relies on personal computing device software but authentication provides some protection against malware threats | Citizen identity relies on secure cryptographic hardware |
| Lack of explicit mutual authentication – eavesdropping, MITM, impersonation of service provider | Citizen cannot easily tell if site or service provider being accessed is genuine | Citizen cannot easily tell if site being accessed is genuine | Citizen can easily tell if site being accessed is genuine |
| Rapidly changing mobile security Rooted / jailbroken devices | Targeted at latest technology | Waits for level of maturity around technology | Security ultimately pinned to secure device distributed to citizen |
| **V.16 Level of Assurance** | | | |
| Low assurance identities susceptible to theft | Low assurance | Medium Assurance | High Assurance |
| Lack of registers containing identity data lowering achievable assurance | Lack of registers | Patchwork of registers of varying quality and completeness | Established authoritative registers |

*Table 7: Vulnerability Types*

### 7.2.2 Scoring

Using the scoring established above we assess the potential severity of vulnerability of the digital identity architectural models. The scores given are only indicative and based on assumptions of how the various approaches work in practice. The specifics of individual schemes may vary.

The goal we believe should be, as far as possible, for schemes to be transparently privacy enhancing. Where design is opaque or reliance is made on the scheme to be operated correctly, this must be assumed to be an element of weakness.

| | Centralised ⟷ Decentralised | | | | | | |
|---|---|---|---|---|---|---|---|
| **Vulnerability assessment** | **Monolithic internet identity provider** | **Federated internet identity providers** | **State issued eID cards** | **Brokered IDPs** | **Brokered CSPs** | **Personal IDP** | **No IDP** |
| **V.1 Fair and lawful processing** | | | | | | | |
| Data held in jurisdictions or organisations with different standards | L | M | L | H | M | L | L |
| Robustness of certification – self certification, privacy focused or not. | H | M | M | M | M | H | H |
| **V,2 Adequacy and quality** | | | | | | | |
| Poor data quality | H | H | L | L | L | M | L |
| Static long lived identities | H | H | M | M | M | H | H |
| **V.3 Explicit and legitimate purposes** | | | | | | | |
| Unnecessary data collection | H | M | L | M | L | L | L |
| **V.4 Minimal disclosure for a constrained use** | | | | | | | |
| Linkable identifiers | H | H | M | M | L | M | L |
| Lack of blinding | H | L | M | L | L | L | L |
| Lack of pseudonymisation | H | H | M | M | L | L | L |
| Lack of "verification mode" services[161] | H | H | M | M | L | M | L |
| Biometric data shared or stored in central database | L | L | H | L | L | L | L |
| **V.5 Openness and transparency** | | | | | | | |
| Lack of transparency | H | H | L | M | L | L | L |
| Implicit data collection | H | H | M | M | M | M | L |
| **V.6 Individual ownership and control of data** | | | | | | | |
| Weak authentication: Authentication token has potential for theft or loss or cloning | M | M | L | M | M | M | L |
| Gap between user and credential (cloud hosted versus card) | H | H | L | H | M | M | M |

[161] i.e. instead of "Is person over 18?" asking "what is person's date of birth"?

| | Centralised ← → Decentralised | | | | | | |
|---|---|---|---|---|---|---|---|
| **Vulnerability assessment** | **Monolithic internet identity provider** | **Federated internet identity providers** | **State issued eID cards** | **Brokered IDPs** | **Brokered CSPs** | **Personal IDP** | **No IDP** |
| Unencrypted data readable without access controls<br>Access to data possible by unauthorised / unauthenticated parties | H | M | M | L | L | L | L |
| Non-expiring user credentials | M | M | L | M | L | M | H |
| **V.7 Accountability and auditing** | | | | | | | |
| Lack of audit | H | H | L | L | L | M | H |
| Access to data not audited | M | M | L | L | L | M | L |
| **V.8 Consent** | | | | | | | |
| Cached consent: Consent once, use many<br>Disclosure of personal information without explicit consent | H | M | L | M | L | L | L |
| No facility to withdraw consent, lack of revocation | H | M | L | L | L | L | L |
| **V.9 Data minimization and avoidance of honeypots** | | | | | | | |
| Pools of data | H | H | L | M | L | L | L |
| **V.10 Sensitive data** | | | | | | | |
| Inferred data (e.g. surname being used to infer religious or ethnic group) | M | M | H | M | L | L | L |
| **V.11 Avoiding exclusion** | | | | | | | |
| High cost to individual or lack of availability | L | L | H | M | M | M | H |
| **V.12 Restrictions on transfer and disclosure of data** | | | | | | | |
| Lack of restrictions | H | H | M | H | L | M | L |
| **V.13 Pluralism and interoperability of systems and technologies** | | | | | | | |
| Proprietary technologies | L | L | M | M | L | L | L |
| **V.14 Minimise the human element** | | | | | | | |
| Reliance on human action (Greater potential for incompetence and negligence to affect system) | L | L | M | M | L | M | L |
| **V.15 Robust technology** | | | | | | | |
| Complexity increasing likelihood of security weaknesses being present | M | M | M | H | M | L | L |
| Susceptibility to malware | H | H | L | M | M | M | M |
| Lack of explicit mutual authentication – eavesdropping, MITM, impersonation of service provider | H | H | M | M | M | M | L |
| Rapidly changing mobile security<br>Rooted / jailbroken devices | H | H | L | H | M | H | H |

| | Centralised ← → Decentralised | | | | | | |
|---|---|---|---|---|---|---|---|
| **Vulnerability assessment** | **Monolithic internet identity provider** | **Federated internet identity providers** | **State issued eID cards** | **Brokered IDPs** | **Brokered CSPs** | **Personal IDP** | **No IDP** |
| **V.16 Level of Assurance** | | | | | | | |
| Low assurance identities susceptible to theft | H | M | L | L | L | L | L |
| Lack of registers lowering assurance achievable | H | H | L | M | M | M | M |

*Table 8: Vulnerability Scoring for the Different Models*

## 7.3 Risks

### 7.3.1 Monolithic Identity Provider

**Examples: Facebook, Google**

| | | |
|---|---|---|
| Key Threats | T.2 | Data breach |
| | T.3 | Individual surveillance |
| | T.4 | Mass surveillance |
| | T.5 | Passing of personal information to unvetted parties |
| | T.7 | Personal information made public |
| Key Areas of Vulnerability | V.2 | Adequacy and quality |
| | V.3 | Explicit and legitimate purposes |
| | V.4 | Minimal disclosure for a constrained use |
| | V.5 | Openness and transparency |
| | V.6 | Individual ownership and control of data |
| | V.7 | Accountability and auditing |
| | V.8 | Consent |
| | V.9 | Data minimization and avoidance of honeypots |
| | V.10 | Sensitive data |
| | V.12 | Restrictions on transfer and disclosure of data |
| | V.15 | Robustness of technology |
| | V.16 | Level of Assurance |

Predictably our analysis reveals numerous threats and potential privacy vulnerabilities with the monolithic Internet identity providers such as Facebook and Google. Their services (and indeed the business models) currently depend on collecting and sharing large volumes of data. Some of the data is obviously personal information (name, date of birth). Other data, including information about the usage of services, can reveal personal details when aggregated, for example, political sites that are 'liked'.

The identity services, especially logon services, revolve around collecting data from the service providers using the identity services. Whilst the user will be prompted to provide consent for their data to be shared, including listing the specific items of data concerned, it is likely that most users will not fully understand what, why and how data is being shared. And where such consent is required in order for the service to be used further, clearly the user will have little option but to consent. Furthermore, both Facebook and Google are known to have shared data with government agencies involved in surveillance.

On Facebook, the Graph and Keyword Insights APIs provide application developers with sophisticated means to interrogate data. In the case of the Graph API only data that is public or where user permission has been given can be accessed. However this depends on the user being aware of and taking time to configure their preferences appropriately and not accepting consent requests. The Keywords Insights API allows statistical information to be returned in response to queries on the use of "keywords" (such as the city of the user posting the keyword). Potentially such queries would allow individual users to be triangulated.

At the time of writing, Facebook have commissioned a consultation with the industry to explore how to achieve "sustainable growth in the data driven economy".[162] As yet it is not clear what the outcome will be and in particular whether it will result in any significant change of direction or approach from Facebook. It is however clear that Facebook are aware of the potential threats to their business of strong privacy regimes and increased customer concerns about privacy including use of ad-blockers.

At face value, the services offered by the Internet giants are by their nature highly inclusive. They are accessible to any connected individual and a great deal of effort is put into the end-user experience, resulting in logon services that are very easy to use. They are not however easy to use in a privacy enhancing manner.

The low assurance of the personal information processed by these platforms also works to some extent in their favour. Currently it probably still is the case that in order to perform identity theft it is necessary to compromise more conventional sources of identity information (intercepting physical mail, phishing bank account and utility data). But as the use of social data for higher value services (including identification services) increases, the risks to such data will also increase.

The primary privacy risks for the monolithic Internet identity providers therefore revolve around the legitimate use of data, following the terms and conditions of those services. Linking social media accounts to higher value identity services, unless done with great care, could expose higher value identity data in ways not intended or desired.

### 7.3.2 Federated Identity Provider

**Examples: GSMA Mobile Connect, Open ID Connect, Mobile Operators, PayPal, Amazon**

| | | |
|---|---|---|
| Key Threats | T.2 | Data breach |
| | T.3 | Individual surveillance |
| | T.4 | Mass surveillance |
| | T.5 | Passing of personal information to unvetted parties |

[162] https://www.ctrl-shift.co.uk/news/2015/10/26/ctrl-shift-and-facebook-exploring-ways-forward-for-the-data-driven-economy/

| | | |
|---|---|---|
| | V.2 | Adequacy and quality |
| | V.3 | Explicit and legitimate purposes |
| | V.4 | Minimal disclosure for a constrained use |
| | V.5 | Openness and transparency |
| | V.6 | Individual ownership and control of data |
| Key Areas of | V.7 | Accountability and auditing |
| Vulnerability | V.8 | Consent |
| | V.9 | Data minimization and avoidance of honeypots |
| | V.10 | Sensitive data |
| | V.12 | Restrictions on transfer and disclosure of data |
| | V.15 | Robustness of technology |
| | V.16 | Level of Assurance |

The risks to federated Internet identity providers will potentially be similar to those of the monolithic identity providers, with two key differences:

- There are potentially a large number of identity providers supporting federated identity protocols such as OpenID Connect. As a consequence each one of them will hold less data about fewer citizens than the monolithic identity providers. Consequently the impact of any breach will be less. Conversely those providers, being smaller, may have fewer resources or less incentive to protect against a breach.
- The commercial models of the federated Internet identity providers may vary. Mobile operators, for example, who are attempting to follow this model with the GSMA Mobile Connect initiative have much less reliance on targeted advertising (although it is an activity they are involved in). Their core business is providing and selling access to mobile network services. As a consequence they may be more able to support digital identity services that are centred around the consumer than an Internet giant that is dependent on targeted advertising as the core of its business.

If, as mentioned in section 6.1.2, GSMA Mobile Connect is linked to the Indian Aadhaar scheme, depending how such a linkage occurs, it could result in data being exposed in unintended ways. Whilst Mobile Connect has defined its own privacy principles[163], mobile operators have often been pursuing propositions involving the monetisation of data.

Separate from Mobile Connect, the Indian government is linking mobile numbers with Aadhaar identities, to enable the mobile channel to be used for alternative forms of authentication, though this has its limitations in view of the tendency of people at the BoP to have up to six SIMs, in order to take advantage of the cheapest mobile phone usage rates available. The mobile phone number is used for example when biometric authentication fails for technical reasons (or the beneficiary doesn't want to do a biometric check), in which case a One Time PIN (OTP) is sent to the beneficiary's mobile phone, for presentment to the service provider. In addition, an eKYC interface exposed by the Unique Identification Authority of India (UIDAI), the government agency charged with the operation of Aadhaar, provides name, address, date of birth, gender, photograph, mobile number and email address simply on the basis of the citizen

[163] https://developer.mobileconnect.io/privacy-principles & http://www.gsma.com/publicpolicy/wp-content/uploads/2012/03/gsmaprivacyprinciples2012.pdf

being able to demonstrate receipt of an SMS.[164] Given the well known issues with SIMs, including poorly-controlled SIM Swap[165], this appears to be a potentially vulnerable service.

### 7.3.3 State Issued eID

**Examples: Estonia eID, Nigeria eID and numerous other national eID schemes – also RealMe, the online-only digital identity issued by the New Zealand Government.**

| | | |
|---|---|---|
| Key Threats | T.1 | Identity theft |
| | T.2 | Data breach |
| | T.4 | Mass surveillance |
| | T.8 | Poor operational security |
| | T.9 | Poor operational processes |
| Key Areas of Vulnerability | V.4 | Minimal disclosure for a constrained use |
| | V.10 | Sensitive data |
| | V.11 | Avoid exclusion |
| | V.12 | Restrictions on transfer and disclosure of data |
| | V.13 | Pluralism and interoperability of systems and technologies |
| | V.14 | Minimise the human element |

State issued eIDs are normally issued in order to provide access to government services. They can also serve as official documents providing access to other services such as KYC for financial services. As a consequence these identities are high value, and could potentially be used to enable fraud if compromised and so become targets for attack. We therefore believe that the key threats to state issued eID are related to identity theft and data breach.

Depending on the objectives of the government and the nature of the implementation, eID systems could allow mass surveillance, especially where there are insufficient firewalls between services or where the identifiers used in the system allow usage to be tracked across multiple services.

The data that is present and contained within eID schemes varies widely. Some, such as Pakistan eID, include religious or ethnic information. Others, such as the Chile eID, do not.

Identifiers may or may not be linkable. Austria is a good example. “Smart” identifiers, where the identifier includes personal information (such as the UK driving licence number which includes part of the citizen’s name and date of birth), clearly enable both disclosure of personal information and linkability.

The majority of state issued eID systems start off with the issuance of a smart card. This is a static technology that does not integrate well with Internet based services. For PC-based access it has been necessary to provide the user with an expensive reader in order to use the smart card.

[164] https://egovstandards.gov.in/sites/default/files/Published%20Documents/Guidelines_on_Mobile_as_Digital_identity.pdf

[165] http://www.bbc.co.uk/programmes/b06w53bh

eIDs are often not integrated as widely into third party services as had been intended. The cost of smart card readers and the cumbersome user experience has often limited the digital use of eID to niche applications.

In our view, the primary privacy risks with state issued eIDs arise from a combination of privacy being insufficiently high on the agenda, and a lack of local expertise in constructing a privacy enhancing system. This can lead to the situation where the eID becomes the container for all manner of data and applications. Understandably some countries (e.g. Nigeria) have copied eID schemes from other countries (e.g. Pakistan); whilst this may promise a more efficient route to implementation, in terms of privacy risks, it may replicate existing risks or fail to adapt properly to the local context.

### 7.3.4 Brokered IDP

**Examples: UK Verify, US Connect.Gov**

| | | |
|---|---|---|
| Key Threats | T.1 | Identity theft |
| | T.2 | Data breach |
| | T.3 | Individual surveillance |
| | T.4 | Mass surveillance |
| | T.5 | Passing of personal information to unvetted parties |
| Key Areas of Vulnerability | V.1 | Fair and lawful processing |
| | V.5 | Openness and transparency |
| | V.6 | Individual ownership and control of data |
| | V.9 | Data minimization and avoidance of honeypots |
| | V.10 | Sensitive data |
| | V.11 | Avoid exclusion |
| | V.12 | Restrictions on transfer and disclosure of data |
| | V.15 | Robustness of technology |

Brokered IDP approaches to identity provide governments with a means to establish a marketplace for digital identity services, potentially sharing the costs of those services with the private sector. At the moment the UK Verify programme is the most advanced scheme[166], although a similar model is being explored in the US (Connect.Gov) and in Australia (DTO).

The threats to these schemes are similar to those for eID. These are identity systems providing access to high value services, including some where the government pays out money, and (if used more widely by the private sector) digital identity for high value services including alternatives to bank KYC. However by creating a marketplace for digital identity it is necessarily the case that IDPs will provide digital services to a much broader range of service providers than just government. Government usage alone will be insufficient to sustain the commercial models of the IDPs. Consequently this raises the possibility that personal information will be shared with organisations that do not have the privacy of personal information as a primary concern unless this use is strictly controlled and enforced. At the moment it is not clear how the

[166] GOV.UK Verify is itself only in public beta with the full public launch anticipated in 2016.

collaboration between public and private sectors will work. For example, at the time of writing the UK Verify programme has just commenced its consultation with the private sector.[167]

Brokered systems are inevitably more complex than single systems. This is especially the case where there are freedoms around the IDP operations. With Verify, for example, the UK defines the technical interfaces to the hub and through its "Good Practice Guides" and "Operations Manual"[168] sets minimum standards with which the IDP must comply. One particular characteristic of the Verify standards is the wide range of documents and sources that can be used for identification. Whilst IDPs are not required to support all of these, the objective is to create a scheme where the individuals entitled to access UK services[169] are able to, including those who do not have access to standard documents such as UK-issued passports and driving licences. Supporting multiple document types inevitably increases the opportunity to a fraudster of being able to find a means to pass the identification process.

Another key feature of the Verify scheme is that citizens are able to create digital identities with multiple IDPs. This is positioned by the UK Government as providing choice and increasing the privacy aspects of the service. For example, a citizen can choose a different IDP (and create separate identities) for HMRC (tax) and DWP (benefits), perhaps in order to keep a separation between interactions with those two Government departments. As there are currently 9 IDPs (in the process of being accredited) and because the registration process can be fully online, there would appear to be a risk that fraudsters may try to set up "parallel" identities with those IDPs not "claimed" by a particular citizen.

Identities in Verify are linkable. The scheme defines identity by a "matching data set" which consists of name, address, date of birth and gender.[170] This data is passed from the IDP, through the hub and to the service provider in every transaction. It is used by the service provider to automatically establish to which customer account the asserted identity relates (a process known as "matching"). Within Verify there is the concept of a service provider-specific (semi) persistent identifier, supposedly with the aim of preventing linkability whilst improving performance. However due to the design it is always accompanied by the matching data set[171].

The presence of a centralised hub in Verify to act as a broker of identity assertions presents an additional risk, which has been the subject of some debate.[172] The UK Government claims that the hub is stateless, storing no data, acting as a "privacy barrier" between IDPs and service providers.[173] There is, however, a need to trust the Government to operate the hub in this way and not to record transaction data, intentionally or otherwise. The architecture is not transparently privacy enhancing; there is a need to trust it.

The issues with Verify arise from specific design choices made for user experience and inclusion reasons. The aim of the UK government is to build a system that works for the entire population (including currently excluded groups) and with minimal friction. The privacy aspects

[167] https://identityassurance.blog.gov.uk/2015/09/30/private-sector-needs-for-identity-assurance-workshop-dates/
[168] https://www.gov.uk/government/collections/identity-assurance-enabling-trusted-transactions#provisioning-identity-assurance
[169] This includes, for example, citizens of other countries who are entitled through marriage to access certain UK benefits such as pension payments. See http://oixuk.org/wp-content/uploads/2015/02/OIX-Existence-Check-Discovery-Project-v1_1.pdf
[170] http://alphagov.github.io/identity-assurance-documentation/shared/glossary.html
[171] http://alphagov.github.io/identity-assurance-documentation/_downloads/GOV.UK_Verify_Architecture_Overview.pdf
[172] http://www0.cs.ucl.ac.uk/staff/G.Danezis/papers/popets15-brokid.pdf
[173] http://alphagov.github.io/identity-assurance-documentation/onboardingGuide/introduction.html

of the hub architecture can be improved (see Section 8.4), but it is likely that there would need to be some trade-off with user experience.

### 7.3.5 Brokered CSP

**Examples: Canada Credential Broker Service (believed to be unique)**

| | | |
|---|---|---|
| Key Threats | T.4 | Mass surveillance |
| Key Areas of Vulnerability | V.1 | Fair and lawful processing |
| | V.6 | Individual ownership and control of data |
| | V.11 | Avoid exclusion |
| | V.15 | Robustness of technology |
| | V.16 | Level of Assurance |

A brokered CSP approach does not provide a solution to the problem of establishing identity in the first place. Rather the focus is on establishing reliable authentication credentials that can be used transactionally to assert an already established identity.

The key issue with this approach is therefore that it does not really help with inclusion. In order to use the Canadian Credential Broker Service the citizen needs to already have a bank account and then be able to link it to an account with the service provider (through some identification process) or create an account (which will involve some separate identification process).

The benefit of this approach is that in mature markets, such as Canada, a system can be created that significantly simplifies accessing services and protects the privacy of the individual. Canada has strong privacy regulations and this approach seems appropriate to that environment.

Whilst no attribute data flows around the system, there is still the potential for identifiers to be linkable, enabling tracking. In the Canadian system, blinding techniques are used to prevent service providers (banks, Government departments) from being able to track usage. Nonetheless, there is still the potential for the hub to link transactions. Furthermore if the scope of the hub is extended to support the exchange of attribute data between service providers, depending on how it is implemented, there is the potential for the privacy characteristics of the hub to be significantly weakened. We believe that this is what has occurred in the US Connect.Gov service (formerly known as FCCX).

### 7.3.6 Personal IDP

**Examples : MyDex, Meeco, Microsoft u-Prove**

| | | |
|---|---|---|
| Key Threats | T.1 | Identity theft |
| | T.3 | Individual surveillance |

| | | |
|---|---|---|
| Key Areas of Vulnerability | V.1 | Fair and lawful processing |
| | V.2 | Adequacy and quality |
| | V.6 | Individual ownership and control of data |
| | V.7 | Accountability and auditing |
| | V.11 | Avoid exclusion |
| | V.12 | Restrictions on transfer and disclosure of data |
| | V.14 | Minimise the human element |
| | V.15 | Robustness of technology |
| | V.16 | Level of Assurance |

Personal IDPs avoid the threat of mass surveillance. By their nature (data encrypted using keys under the control of the individual) it appears that they are much less susceptible to scalable data breaches. Properly implemented, data stored in cloud servers will always be encrypted and hence any breach would not compromise that data. Depending on how the data is organised (e.g. if there is structure to the data), it may be possible to infer information (e.g. that an individual has certain attributes and not others), but the data itself should be protected.

A key aspect of personal IDPs is that greater control is passed to the individual. They promise citizens more granular levels of consent, but this can potentially place a significant burden on the citizen in the maintenance of data and in having some discretion regarding how and where data is shared.

Depending on the implementation, Personal IDPs place a greater reliance on the security of the devices the citizen uses to manage their identity. Even where data is stored in a cloud service, ultimately the security of that data will boil down to how the citizen's private cryptographic keys are protected. If these are held in software in a personal computing device (either a PC or phone), then they are at risk of theft (via malware) or loss (if no backup arrangement is in place). The security of mobile devices in particular is rapidly changing. The security controls employed in mobile operating systems are generally well designed and provide relatively safe sandboxed environments in which applications can run. However, mobile operating systems are upgraded frequently and there is a great deal of effort in the hacker community to find new ways to work around security controls (especially to find ways to root or jail break devices and to identify other exploits).

The privacy risks in personal IDPs should be lower than more centralised services. However this places greater demands on the individual, including computer proficiency and self sufficiency, to manage their own personal information.

### 7.3.7 No IDP

**Examples: Bitcoin, various Blockchain identity startups such as Shocard**

| | | |
|---|---|---|
| Key Threats | T.1 | Identity theft |
| | T.3 | Individual surveillance |
| | T.4 | Mass surveillance |

| | | |
|---|---|---|
| Key Areas of Vulnerability | V.1 | Fair and lawful processing |
| | V.2 | Adequacy and quality |
| | V.6 | Individual ownership and control of data |
| | V.7 | Accountability and auditing |
| | V.11 | Avoid exclusion |
| | V.12 | Restrictions on transfer and disclosure of data |
| | V.14 | Minimise the human element |
| | V.15 | Robustness of technology |
| | V.16 | Level of Assurance |

Bitcoin was created to provide a digital currency that could operate without regulatory oversight, without the need for banks to act as intermediaries and to avoid censorship. Whilst this presents challenges to governments it has resulted in security architectures that appear to be highly robust, in terms of privacy, security and resilience.

The approach to identity in Bitcoin is that users create and manage their own cryptographic keys, which are used to establish ownership of the currency. No other data need be attached to the keys and in this sense identities can be viewed as anonymous. In reality it is possible sometimes to infer some information about the owners of particular cryptographic identities by analysing linkages between transactions (which are all openly readable on the Bitcoin blockchain) as well as the meta-data associated with the transactions.

The technological concepts and architecture of Bitcoin have developed into what is now typically referred to as either “blockchain technology” or “distributed shared ledger technology”. At the moment these are quite loose concepts but typically include:

- A ledger (a record of the history of ownership) that is replicated across the nodes of the network.
- A consensus mechanism that ensures the nodes agree on which is the correct version of the ledger.
- Strong cryptographic security ensuring that records are immutable.
- Increasingly a focus on smart contracts to automatically enforce and potentially automatically execute business rules relating to the assets.
- The potential for strong levels of privacy through cryptographic blinding techniques such as “Zero Knowledge Proof”.

From an identity perspective these technologies have the following characteristics:

- They are very immature and no clear model for digital identity has been established.
- For “public” blockchains (such as Bitcoin) where there is no clear governance, upgrading the network (for example to support longer cryptographic key lengths over time[174]) will be difficult. This will be less of an issue for “private” blockchains.
- The onus is on the end user to manage their identity, i.e. there is “no IDP”.

[174] This is a basic requirement of any cryptographic service.

Many citizens will not have the ability to create and manage cryptographic identities. It is therefore necessary for key management services to be provided, for example, by wallet providers. This has been one of the main causes of loss to users in Bitcoin. The famous Mt Gox breach did not undermine the Bitcoin network, per se, but rather compromised the digital identities of the individuals who had placed their trust in Mt Gox to manage their identities.

In the future, Blockchain technology may be able to offer high levels of both privacy and transparency (or accountability), which are currently business requirements that need to be traded off against each other. For the moment it remains a catalyst of innovation, with the true potential yet to be determined.

## 7.4 Privacy trade offs

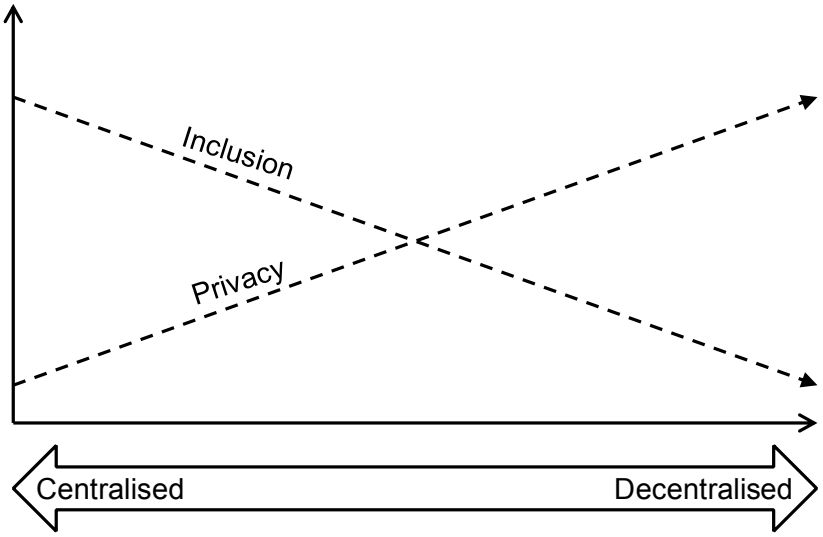



*Figure 18: Privacy Trade offs*

The analysis provided in this chapter suggests there is often a trade off between privacy and inclusion. Sometimes this is a deliberate choice on the part of the identity provider. At other times it is an implicit consequence of the business model or architecture adopted by the identity provider (or scheme to which they belong).

At one extreme the highly centralised monolithic Internet identity providers have commercial incentives to build inclusive services that depend on free, simple and ubiquitous access to citizens. This has enabled them to build billion dollar advertising businesses. However this has come at the cost of the privacy of the personal information of citizens. Whilst these organisations are reacting to demands to address privacy concerns, it is far from clear if and how their business models will adapt.

At the other extreme, highly decentralised and consumer-centric services offer high levels of privacy. These however suffer from issues related to inclusion. Firstly there is a greater burden on the consumer to manage their own personal data and digital identity credentials. Organisations offering services in this space are continually seeking to refine the way that consents and permissions are managed, and it is not clear if there is an optimal balance between the granularity (or complexity) of control available to the citizen, and the simplicity of use. Secondly there is often a “chicken and egg” problem with decentralised digital identity

services; service providers will only integrate digital identity services that enable a large number of customers, but customers will only sign up to digital identity services that have utility at enough service providers. Last but not least, the business models around decentralised digital identity services lack maturity.

Between these two extremes are models that offer some level of privacy and as well as some level of inclusion. The foundational national level schemes (including national eID schemes as well as the brokered digital systems) described in this document sit in this middle ground. As illustrated above whilst current systems often have privacy issues, these generally arise from specific design choices. Consequently with good design we believe it should be possible to achieve an acceptable balance between privacy, inclusion and other business requirements such as user experience (or reducing friction in the use of services). It certainly is possible to learn from current experience to make significant improvements to existing systems to produce architectures and designs that are more transparently privacy enhancing, and reduce the amount of trust that must be placed in the operation of the digital identity service.

Informed choices are needed about these important issues and this needs the development of suitable capacities to understand and agree them. Such capacities are not necessarily found amongst politicians but should also not be delegated to industry (or academics)[175]

[175] Whitley, E. A., and Hosein, I. R. 2008. “Doing the politics of technological decision making: Due process and the debate about identity cards in the UK,” European Journal of Information Systems (17:6), pp. 668–677.

# 8 MITIGATION

Having determined that there are numerous potential risks to privacy and the privacy of personal information, this chapter considers some of the mitigation strategies that might be employed.

From the outset the following observations are made:

- The parties that are required to mitigate risk may not be the parties with the interest in doing so. The internet giants again provide a clear example of this. Their primary interest is in maintaining advertising revenues, which come from advertisers not consumers.
- Complex innovative systems can develop in unpredictable ways with unintended consequences. This is especially true where there is pressure to roll out new features and service at a rapid pace, meaning that there may be little time for a measured consideration of potential privacy issues.

This chapter outlines some of the ways in which privacy risks in digital identity systems could be mitigated. Whether there is sufficient motivation to undertake such measures will depend on a combination of political, social and commercial pressures.

## 8.1 What are we trying to mitigate?

As explained in section 4.7, trust frameworks are the identity industry mechanism for establishing the scope, purpose and qualities of a particular digital identity service. These provide the "rules" that govern the way in which the service is operated and usually fall into three areas: regulatory, technical and commercial.

It should therefore be expected that privacy issues can arise in or be caused by each of these areas. This is demonstrated in sections 5 and 7, which consider some of the specific privacy impacts arising from different digital identity models and the potential risks in those models.

The following table summarises the key areas of impact and risk:

| | |
|---|---|
| Regulatory | • Inadequate data protection law<br>• Weak institutional mechanisms |
| Technical | • Identifiers that are linkable<br>• Unnecessary collection of data<br>• Unintended disclosure of data |
| Commercial | • Commercial models, especially those based on deriving additional value from user data, in conflict with privacy. |

*Table 9: Key areas of impact and risk*

## 8.2 Where are we starting from?

Perhaps the most important factor determining what mitigation strategy is appropriate or feasible will be the status of the digital identity system in question.

### 8.2.1 Greenfield

If a digital identity system is to be developed from scratch, that may provide the opportunity to "design in" privacy from the beginning. However there are still many factors that will influence the overall approach taken to digital identity and which will as a result place constraints on the mitigation strategies that can be employed.

Digital identity systems are a means to an end and not an end in themselves. They enable other digital services. For a digital identity system to be successful it must address the needs of the services it is to enable. Often the business-led requirements of digital services are focused on areas other than privacy, such as lowering friction, building scale and ensuring a positive business case. This is certainly the case for private sector led schemes, and may also apply to state-sponsored schemes. For example, cost was a key factor behind the UK Government's decision to pursue a brokered model, and development of actor-led schemes in areas such as voter registration and health services are likely to be similarly constrained.

There may be a desire to copy an existing digital identity system (a "blueprint"). This may provide a fast route to implementation and allow some sharing of costs but it will bring potential issues. The existing digital identity system may have been created to address specific market conditions (e.g. attitudes to privacy, structure of government) that do not apply in the new context. Any reuse of an existing system must involve a proper privacy impact assessment in order to ensure that the privacy needs in the new context are addressed.

There may be political or social pressure to take certain approaches. A clear example of this is the UK's experience, where the No2ID campaign and the LSE report[176] in the UK were instrumental in the demise of the UK Identity Card[177][178][179], due to resistance to the idea of a compulsory identity card in general, and the proposed creation of a centralised 'honeypot' of identity and biometric data in particular. This was another key driver for the current brokered model followed by Verify in the UK.

### 8.2.2 Regional Influence

State Issued eID cards in particular may need to interoperate with the infrastructures of neighbouring territories. In the GCC countries for example eID cards can be used as travel documents across the region.[180] In Europe the recently passed eIDAS regulation is focused on enabling cross border acceptance of European issued eIDs in both face-to-face and online contexts.[181]

Clearly where such interoperability is required this may constrain the privacy of the digital identity system, if for example some minimum data is required for transactions to be

[176] identityproject.lse.ac.uk
[177] http://researchbriefings.parliament.uk/ResearchBriefing/Summary/LLN-2016-0002
[178] http://www.publications.parliament.uk/pa/cm201011/cmhansrd/cm100609/debtext/100609-0009.htm#10060953001109
[179] https://www.eff.org/pages/success-story-dismantling-uk's-biometric-id-database
[180] http://www.arabnews.com/id-cards-let-gcc-nationals-breeze-through-airports
[181] http://ec.europa.eu/digital-agenda/en/trust-services-and-eid

interoperable or if interoperability forces other design constraints on the system. In the case of travel interoperability, the minimum required data is specified by ICAO.

Enabling interoperability between digital identity systems can introduce new risks. In the eIDAS regulation two implementation models are supported: centralised and decentralised. One member state may choose to implement a decentralised model for privacy reasons but when the eIDs it issued are used to access cross border services in other member states who have adopted a centralised model, that other member state will be able to track usage across services located in its jurisdiction.[182]

### 8.2.3 Existing Legacy System

Where digital identity systems already exist addressing privacy issues may be difficult, due to the cost of making changes and the potential impacts to systems. This will be particularly true if those digital identity systems are integrated into many services and applications. For example, if a service collects all available digital identity data and then a change is made to restrict the data that is shared, this could directly impact the service in question.

This does not mean that privacy improvements cannot be made. As an example of this, contactless EMV[183] payments cards originally included the cardholder name as a field that could be read over the contactless interface. Recognising that this presented a privacy issue, due to the potential to track individual real names, the payment industry now requires that field to be blanked out.

### 8.2.4 The stages of development

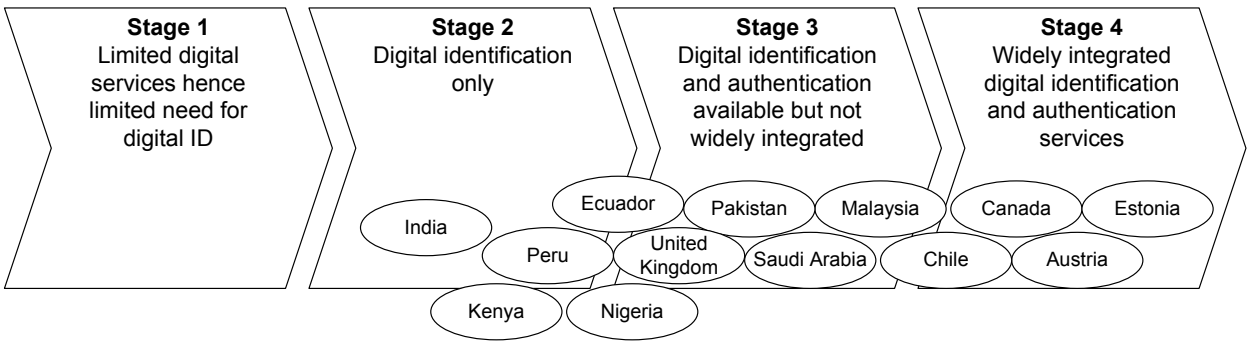



*Figure 19: Stages of development*

Figure 19 illustrates at a high level the stages that a country may go through in the development of a foundational nation-scale digital identity and overlays approximately where the countries considered in this report fit on this scale. The following observations are made:

- The Aadhaar programme in India has been primarily focused on establishing a register of biometric data. Various initiatives are underway to link authentication credentials to that data in order to create usable digital identities but for now these are not in widespread usage.

[182] https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/ElekAusweise/SmartCard_Workshop/Workshop_2015_Bender.pdf?__blob=publicationFile

[183] https://www.emvco.com/

- There are several countries such as Peru, Nigeria, Ecuador and Kenya which have commenced eID programmes but for which the number of eID cards issued is a relatively small percentage of the population or nil. In writing this report, we were informed by one of these countries that the development of future services and wider integration of the eID is predicated on a wider take up of the eID.
- Once a digital identity scheme has been established it will not provide real utility to citizens until it is integrated into a wide range of services. There are several countries (e.g. Pakistan, Malaysia, Saudi Arabia) which appear to have digital systems integrated into digital services but information is not readily available concerning the number of services or users. In Figure 19, stage 3 represents this stage where the identity scheme has been delivered but is yet to be supported by a wide range of digital services. Stage 4 on the other hand represents the mature state where the digital identity scheme is widely integrated and the de facto means of asserting identity to government in a digital context.

From the point of view of mitigating privacy impacts, it is going to be simpler to implement changes or influence change in countries in earlier stages of development. These will have fewer services dependent on the digital identity working in a particular way.

In this report we focus on reviewing the status of digital identity programmes. There will be many examples of digital services that use a form of identification (e.g. entering a social security number into a web site) but without authentication. Such services, especially if they are high value, would clearly benefit from digital identity (including digital identification, authentication and authorisation) to both reduce fraud and if implemented appropriately to enhance privacy.

## 8.3 The role of privacy of information principles in mitigation

Privacy of information principles are widely recognised as being an important way of guiding how privacy is approached. They help to ensure privacy is considered in a broad sense and inform good practice. However in themselves they do not guarantee that the privacy of information is appropriately addressed.

Because principles are necessarily broad in nature, they will inevitably be open to some level of interpretation. For example, a merchant may have a different view on whether data about the age of a customer is “relevant” to a purchase, compared to the customer.

As with all rules and guidelines, a provider (of identity or services) could adopt a checkbox mentality to the principles.

Lack of appropriate skills or relevant experience may result in systems being developed which superficially follow the principles but in reality contain significant gaps.

As with information security, privacy is a moving target. The risks to the privacy of personal information will change over time with the evolving threat landscape and constantly changing technology. The privacy of information is therefore not something that should be considered only during the implementation of a service. It needs to be constantly reviewed and monitored.

The management of privacy risks should be owned and be the responsibility of the most senior stakeholders in an organisation (or government). They should ensure that the importance of

privacy is set through appropriate policy and the structures are in place to implement those policies in appropriate ways.

Further consideration of privacy principles for digital identity is set out in Section 9, where we also set out our digital identity privacy principles (DIPPs).

## 8.4 Mitigation of risks: Specifics for each model

### 8.4.1 Monolithic Identity Provider

**Examples: Facebook, Google**

| | |
|---|---|
| **Primary risk**: | Mass sharing of personal information with unaccountable or malicious third parties |

The current business model of the monolithic identity providers relies on collecting and sharing data. Consumers have also been conditioned to expect access to those services for free.

These organisations have been making incremental changes to the privacy features provided to consumers. Typically these have involved increasing the granularity of control users have in what information they share with other users. They have not typically altered the way that data is shared with third party apps nor changed the fundamental business model on which the services are based.

Mitigating the risks with these identity providers will potentially require a significant change in business model.

One can envisage models where the monolithic IDPs address target higher value digital identity services by establishing digital identity databases that are separate from their core internet services. This could be similar to the model already employed by the credit reference agencies who already maintain databases of credit data (governed by the rules of reciprocity) that are segregated from their broader data aggregation business. They bring the same data analytics skills to bear on both data set types. This type of approach may allow the monolithic IDPs to offer increased privacy for higher value digital identity services. Whether such an approach, and divergence from their core business model, would be of sufficient interest to the monolithic IDPs is unclear.

### 8.4.2 Federated Internet Identity Provider

**Examples: GSMA Mobile Connect, Open ID Connect, Mobile Operators, PayPal, Amazon**

| | |
|---|---|
| **Primary risk**: | Some sharing of personal information with unaccountable or malicious third parties |

The Federated Internet Identity Provider model is similar to the Monolithic Internet Identity Provider model except that identity providers may have alternative business models that are less dependent on them profiling users by acquiring and monetising personal data.

Having said that, knowing and understanding the customer is critical to any business and for consumer facing digital businesses this has usually entailed establishing some type of CRM[184] capability where customer data and service usage data is collected, analysed and leveraged.

To mitigate the risk of sharing personal information, an organisation that knows the customer (e.g. a mobile operator) and has a genuine concern for privacy establishes an identity provider service that provides specific attribute assertions (e.g. age, location) in controlled ways. The issue here is likely to be how the consumer can, in an open market, determine which identity providers have a greater interest in privacy. In addition, unless constrained by contract or regulation, there may be no guarantees that the motivation and practices of the identity provider will not change.

Federated internet identity providers of any size would be a target for data attacks, given that without additional layers of obfuscation and blinding it would have to be assumed that the provider would hold large pools of data.

Federation protocols, such as OpenID Connect, allow the identity provider to see which service provider is being accessed by the user. By modifying transaction flows and protocols it may be possible to address this issue.[185] In some cases the end result is the Personal IDP model described in this document. In others, new models or technologies would need to be developed that are not currently available in the market.

### 8.4.3 State Issued eID

**Examples: Estonia eID, Nigeria eID and numerous other national eID schemes – also RealMe, the online-only digital identity issued by the New Zealand Government.**

| | |
|---|---|
| **Primary risk**: | Insufficient attention given to privacy in design, implementation and operation |

The Austrian eID is often held up as an exemplar of a privacy-enhancing state issued eID.[186] It is the result of a political and regulatory environment sensitive to the issues of privacy, and due to a carefully constructed privacy-led design.

Where states have implemented eID systems that are weaker in terms of privacy this could be due to the political and regulatory environment being insufficiently focused on privacy to demand a privacy enhancing approach at the time of development or renewal. It could also be due to a lack of privacy expertise, both in setting out the principles but also in following those principles through into the architecture and design.

For state issued eID systems, there is perhaps a need for guidelines to assist authorities to illustrate good technical approaches to addressing privacy principles. These could include items such as:

- Hold the minimum amount of data possible in the eID system. In the best case just an unlinkable identifier that conveys no personal information but which can be linked by the user to the relevant user account at the service provider.

[184] Customer Relationship Management
[185] http://arxiv.org/ftp/arxiv/papers/1401/1401.4726.pdf
[186] http://ec.europa.eu/idabc/en/document/4486/5584.html

- Use separate identifiers for each service access by the user to prevent tracking of usage.
- Use strong cryptographic hardware (e.g. smart card, SIM in mobile) to protect user credentials. Note that alternative mobile technology such as Trusted Execution Environments, white box cryptography and software hardening may be able to provide a more cost effective way to achieve an acceptable level of protection for user credentials.
- Ensure that the data held by the individual service providers leveraging the eID is minimised, so that each provider only holds the data they genuinely need.
- Ensure there is no communication or sharing of data between service providers enabled by the eID system.

### 8.4.4 Brokered IDP

**Examples: UK Verify, US Connect.Gov**

| | |
|---|---|
| **Primary risk**: | Hub architecture provides point of aggregation |

Hub architectures have been adopted as they are a known and proven way to establish market infrastructure. The payments industry, for example, is currently dominated by hub architectures, which are sometimes also referred to as switches, when operated by the likes of Visa and MasterCard. These hubs were created pre-internet when consumers, merchants and banks were not connected to each other. They provided the network (albeit hub and spoke in architecture). With the ubiquitous connectivity provided by the internet, hubs should not be necessary, although there might still be a need for directory or discovery services.

The issue with hubs is that whilst they may claim to be stateless (i.e. have no requirement to cache or store the business or personal data that flows through them) and hide the relying party from the identity provider, they do provide a point of aggregation and if compromised could start to build a store of personal data, in particular tracking activity across the hub.

To mitigate these issues and to encourage hubs to be more transparently privacy enhancing, we believe the following types of measure should be employed:

- Use different unlinkable identifiers for different relationships established across the hub.
- Encrypt personal data end-to-end (i.e. from sender to receiver with no break in encryption at the hub), so that it is inaccessible to the operator of the hub.

The aim is to reduce the hub to being simply a communications broker providing interoperability benefits whilst minimising the potential access to personal information.

Mitigating these issues at the business application level (e.g. by using cryptographic blinding techniques) may be insufficient to prevent tracking altogether. Hubs may still be able to track at a lower level in the communications stack. For example, depending on the transaction flow and session control, it may be possible for the hub to determine the IP address and even the physical location of the citizen and link this with the various service providers accessed and

identity providers used. It is unclear whether these low level issues can be avoided in a centralised hub architecture.

The alternative to a centralised hub is a decentralised hub, where each citizen has their own personal hub running locally on their own device, brokering the identity transactions such that the only party that can track the citizen is the citizen themselves. This is covered by the personal IDP model considered in this document.

### 8.4.5 Brokered CSP

**Examples: Canada Credential Broker Service (believed to be unique)**

| | |
|---|---|
| **Primary risk**: | Hub architecture provides point of aggregation |

The issues and hence the mitigation strategies for Brokered CSPs are the same as for Brokered IDPs. Brokered CSPs start from a better privacy position, as they are brokering authentication actions only. This architecture still employs a hub which can act as a point of aggregation and therefore the same mitigation strategies should be employed as with Brokered ISPs.

### 8.4.6 Personal IDP

**Examples : MyDex, Meeco, Microsoft u-Prove**

| | |
|---|---|
| **Primary risk**: | Individual unable to manage personal information effectively |

Consumer centric digital identity, whether in the form of a personal data store or claims-based identity technology, is designed with privacy in mind. In it the user is made the arbiter of their personal data usage. They allow attribute assertions to be collected (or sourced) from attribute providers without the attribute provider ever being able to tell where the attribute assertion is to be used. This prevents tracking and puts the user in control, determining when and where attribute assertions will be shared.

Of course to have any value an attribute assertion needs to come from a trusted source (such as an authoritative government registry or a trusted private sector organisation such as a bank) and be verifiable against that source. This would typically mean that the attribute is signed by that source and the signature can be verified using a public key certificate from the source. Such an approach could reveal information about the source (e.g. who the consumer banks with or which mobile operator they use) to the service provider consuming the attribute assertion. With the Attribute-Based Credential technology researched by the EU's ABC4Trust programme[187] the credential issued to the individual is not directly provided to the service provider. Instead a "presentation token" is used which is derived from the issued credential which provides the opportunity for additional blinding, including potentially of the issuer.[188]

Personal IDPs therefore have the potential to provide robust technical information privacy. There would however appear to be a greater reliance for the citizen to manage the data within

[187] https://abc4trust.eu/download/ABC4Trust-D2.1-Architecture-V1.2.pdf
[188] E.g. a service provider could agree to accept credentials from all banks certified to a particular standard and know that the presentation token is derived from a credential issued by one of the certified banks but not be able to tell which one.

their personal IDP. If attributes are to be sourced from a wide range of locations this has the potential to be complex for the user. In a mature market with widely adopted standards and established user experience this would be less of an issue but this is not the current state of the market. A more pragmatic approach therefore may be to limit the attributes and sources in order to allow a basic level of identity services to be established.

### 8.4.7 No IDP

**Examples: Bitcoin, various Blockchain identity startups such as Shocard**

| | |
|---|---|
| **Primary risk**: | Technology promises high degree of privacy but is unproven |

Privacy is a key focus of the Blockchain community. The means through which privacy could be achieved has a large overlap with Attribute-Based Credential technology. Both use cryptographic techniques, such as Zero Knowledge Proof, to ensure that the absolute minimum data is shared in any interaction.

The hype around Blockchain has resulted in it becoming a byword for innovation. Numerous organisations are experimenting with various flavours of the technology. Some of these experiments have an identity focus. It is too early to know exactly what the relationship between digital identity and blockchain technology will be. What Blockchain does is raise the bar for what is viewed as privacy enhancing. In particular, the starting point for digital identity in blockchain is an unlinkable secure identity (a cryptographic public/private key pair) that is only linked to transactions, digital assets or other data with explicit user action. There is also no restriction on the number of identities an individual may choose to create, potentially allowing the user to prevent unwanted linking of transactions.

## 8.5 Summary

Every model for digital identity considered in this document has the potential to be poorly implemented from the point of view of privacy. A more decentralised or user-centric design however encourages the system to be focused around the individual and we believe will ultimately result in an approach that is more transparently private.

Whenever a digital identity system is more centralised, there is a need to trust the operator of centralised services. Even when those services are operated by an organisation such as a bank or government, that a citizen might reasonably trust, data breaches can still occur.

There are however some basic steps that will help to ensure the model adopted promotes a good approach to privacy, including:

- Making the individual the control point for their digital identity. Ideally this should be achieved by encrypting attributes with cryptographic keys only normally[189] accessible to the individual.
- Use unlinkable identifiers to prevent the matching of individuals across their use of services.

[189] Careful consideration need will also need to be given to how cryptographic keys or attribute data is recovered when the individual loses their keys.

- Building end-to-end encryption into transactions to avoid unintended leaking of data.
- Use a trust framework to provide an explicit, open and transparent description of how the digital identity service works in order to be open to public scrutiny as well as providing a means for certification.

Whichever approach is taken the privacy requirements will need to be balanced against other requirements such as usability, inclusion. The most private solution may not always be the most practical solution. These potential trade-offs are explored in Section11.

# 9 PRINCIPLES

## 9.1 Overview of privacy principles

It is broadly accepted that in order to safeguard personal data, and thus the right to informational privacy, it is necessary to promote regulation relating to the collection, use, storage and destruction of such data. Data protection regulation, which forms part of the legislative frameworks of more than 100 countries around the world, has its origins in the 1980 OECD Guidelines on the protection of personal data. These Guidelines articulated eight principles of informational privacy, which closely echo the Fair Information Principles enshrined in US law and policy, and which have come to heavily influence the development of data protection regulation around the world.

| Principles contained in the 1980 OECD Guidelines | |
|---|---|
| Collection Limitation | There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject. |
| Data Quality | Personal data should be relevant to the purposes for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up-to-date. |
| Purpose Specification | The purposes for which personal data are collected should be specified at the time of collection, and any subsequent use should be limited to the fulfilment of those purposes or others as are not incompatible with those purposes; |
| Use Limitation | Personal data should not be disclosed, made available or otherwise used for purposes other than those specified in accordance with the purpose specification principle, except with the consent of the data subject or by the authority of the law. |
| Security Safeguards | Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorised access, destruction, use, modification or disclosure of data. |
| Openness | There should be a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity and usual residence of the data controller. |
| Individual Participation | Individuals should have the right to obtain data from a data controller in a reasonable manner, form and at a charge. |
| Accountability | A data controller should be accountable for complying with measures which give effect to the principles stated above. |

*Table 10: OECD Guidelines*

From the early 1990s, the European Commission developed the content of the European Data Protection Directive (DPD), ultimately adopted in 1995. The DPD was based on the OECD Guidelines but sought to advance the content and scope of the principles, and the strength of their application. While the challenges posed by technology were primary drivers of the Directive (as evidenced by the focus on "automated" processing of personal data) the DPD was still negotiated in a pre-Internet Europe, and thus lacks application and relevance to many current-day data protection contexts.

In addition to elements which correspond to the eight OECD Principles, the European DPD includes a number of other areas of regulation, including:

- Fair and lawful processing - A general requirement of 'fair and lawful processing', beyond collection limitation, was introduced
- Deletion - Destruction or anonymisation of personal data after the purposes for which it is held are completed
- Sensitive data protections - Additional protections for particular categories of sensitive data
- Automated processing controls - Data controllers should ensure that automated decision-making which significantly affects data subject is subject to human checking, and data subjects should be able to know the logic of such automated data processing
- Data export restrictions based on destination - Requirement of restrictions on personal data exports to countries which did not have a sufficient standard of privacy protection

The European DPD principles remain, internationally, the state of the art, although the EU is currently negotiating a new Regulation which, when adopted, will replace and update the DPD. The DPD applies to private and public actors across all sectors and situations, including in the context of identity systems.

In addition, a further set of Principles has since been developed for application to the particular field of digital identity. Canadian computer scientist, and Microsoft's Chief Architect of Access, Kim Cameron published in 2005 a thesis entitled the *7 Laws of Identity,*[190] which seeks to set out a number of principles applicable to the particular dynamics of digital identity systems, and assist in creating "an identity layer for the internet".

Cameron's Laws of Identity enshrine some of the OECD and European DPD concepts, such as user control and consent, and minimal disclosure of information, yet they do so with a particular view to digital identity and, in particular, the need to adapt to a range of technologies and services on the internet. The seven Laws are, in brief:

1. Technical identity systems must only reveal information identifying a user with the user's consent.
2. The solution which discloses the least amount of identifying information and best limits its use is the most stable long term solution.

[190] Available at http://www.identityblog.com/stories/2005/05/13/TheLawsOfIdentity.pdf

3. Digital identity systems must be designed so the disclosure of identifying information is limited to parties having a necessary and justifiable place in a given identity relationship.

4. A universal identity system must support both “omni-directional” identifiers for use by public entities and “unidirectional” identifiers for use by private entities, thus facilitating discovery while preventing unnecessary release of correlation handles.

5. A universal identity system must channel and enable the inter-working of multiple identity technologies run by multiple identity providers.

6. The universal identity meta-system must define the human user to be a component of the distributed system integrated through unambiguous human-machine communication mechanisms offering protection against identity attacks.

7. The unifying identity meta-system must guarantee its users a simple, consistent experience while enabling separation of contexts through multiple operators and technologies.

Cameron’s Laws are focused more on the usability and effectiveness of digital identity systems rather than on the protection of information from misuse and abuse per se, and the avoidance of data stores vulnerable to attack.

In that regard, the UK government’s Verify identity programme has developed its own set of principles, the Identity Assurance Principles, which purport to address informational privacy issues while at the same time being tailored to the specific challenges and opportunities arising in digital identity schemes. Again, the principles take as their starting point data protection principles, but have particular elements specifically applicable to the Verify programme, such as the Governance/Certification Principle, which relates to individual trust dependent upon the certification of private sector identity providers by the government.

Determining the extent to which digital identity schemes impact upon individual privacy rights is facilitated through the utilization of transferable principles against which the specificities of respective digital identity initiatives can be objectively compared. Although the OECD Principles provide the strongest orientation to informational privacy, they are high-level and can benefit from elaboration in the context of digital identity; on the other hand, Cameron’s Laws of Identity are perhaps too oriented towards practical barriers and do not cover many of the elements that could make a digital identity scheme reinforce informational privacy rights.

The various sets of privacy principles are compared in Table 11 which illustrates the fragmented nature of this field.

| Privacy in Identity Principles | Mapping to existing sets of privacy or identity principles | | | | |
|---|---|---|---|---|---|
| | OECD Principles | Fair Information Practice Principles | EU Data Protection Directive | Laws of Identity | Verify IAPs |
| Fair and lawful processing | Collection limitation | | Data to be processed fairly and lawfully | | |
| Adequacy and quality | Data quality | Data minimisation<br>Data quality and integrity | Data to be adequate, relevant and non excessive; accurate and up to date | Minimal disclosure for a constrained use | Data quality |
| Explicit and legitimate purposes | Purpose specification | Purpose specification | Data to be collected and processed for specified, explicit and legitimate purposes | Justifiable parties | |
| Minimal disclosure for a constrained use | Use limitation | Use limitation | Data not to be further processed in way incompatible with specified purposes | Minimal disclosure for a constrained use | |
| Ensure highest levels of security that integrate the individual | Security safeguards | Security | Appropriate technical and organisational measures | Human integration | |
| Openness and transparency | Openness | Transparency | Information to be given to data subject | | Transparency |
| Individual ownership and control of data | Individual participation | Individual participation | Right of access to and correction of data | User control and consent | Service-user access and portability |
| Accountability and auditing | Accountability | Accountability and Auditing | Data controllers responsible for implementation | | Problem resolution |

| Privacy in Identity Principles | Mapping to existing sets of privacy or identity principles | | | | |
|---|---|---|---|---|---|
| | OECD Principles | Fair Information Practice Principles | EU Data Protection Directive | Laws of Identity | Verify IAPs |
| Consent | | Individual participation (seek consent "to the extent practicable") | Data to only be processed with consent (with exceptions) | User control and consent | User control |
| Data minimization and avoidance of honey pots | | Data minimisation | Data minimisation | Minimal disclosure for a constrained use | Data minimisation |
| | | | Deletion or anonymisation of data after use | | |
| Sensitive data | | | Sensitive data protections | | |
| Avoid exclusion | | | Automated processing controls | | |
| Restrictions on transfer and disclosure of data | | | Restrictions on transfer of data | | |
| | | | | Directed Identity | |
| Pluralism and interoperability of systems and technologies | | | | Pluralism of operators and technologies | Multiplicity |
| | | | | Consistent experience across contexts | |
| | | | | | Governance certification |
| | | | | | Exceptional circumstances |

*Table 11: Comparison of Privacy Principles*

## 9.2 Digital identity privacy principles

The fragmented nature of the existing principle sets in this field means that there is no one set of principles that meets all our needs, i.e. that covers legal considerations, technical measures and commercial issues. So for the purpose of this paper we have developed our own set of principles, the Digital Identity Privacy Principles (DIPPs), which encapsulate the major relevant principles from the other five sets of Principles discussed here. The *Digital Identity Privacy Principles* are as follows:

**Fairly Regulated**

1. Fair and lawful processing – Identity schemes must be regulated by strong legal frameworks that require data to be processed fairly and lawfully
2. Adequacy and quality – Data is to be adequate, relevant and non-excessive in relation to the purpose for which it was collected, accurate and up to date
3. Explicit and legitimate purposes – Data is only to be collected and processed for specified, explicit and legitimate purposes
4. Minimal disclosure for a constrained use – Data is only to be used for ways which are compatible with specified purposes and disclosed to parties only to the extent it is strictly necessary

**Citizen Centric**

5. Openness and transparency – Individuals should be given the greatest amount of information possible about how their data is processed, used, stored, disclosed, retained and deleted
6. Individual ownership and control of data – individuals should have the ability to access information about what data is held about them, who has access to that data and on what conditions it is being processed. Individuals should be able to correct and transfer their data where applicable, as well as rights to stop uses and have data deleted.
7. Accountability and auditing – there should be regular independent auditing of identity providers, and individuals should have avenues of redress if their data is misused or incorrectly disclosed. Additional protections (such as anonymization / deletion of older audit trails) should be considered.
8. Consent – identity schemes should always have the consent of individuals to use their data for the purposes specified.

**Protecting privacy**

9. Data minimization and avoidance of honeypots – Identity providers and others involved in the identity process should request and store the minimal amount of data necessary to perform their functions, thus minimising the creation of data honeypots at all times.

10. Sensitive data – there should be specific protections for sensitive data (that relating to political, religious, ethnic identity or health data).

11. Avoiding exclusion – identity providers have an obligation to ensure that individuals' data is not used in a way that excludes them from access to services and opportunities they are entitled to.

12. Restrictions on international transfer– data should not be transferred to parties or locations outside of the data controller's direct control.

**Interoperable and secure**

13. Pluralism and interoperability of systems and technologies – identity providers should choose systems and technologies that are useable, transferable and interoperable, and incorporate the individual.

14. Minimise the human element – The more a system is reliant on human intervention, the greater the potential for negligence and abuse.

15. Robust technology – The system incorporates strong cryptographic software and hardware.

16. Levels of assurance – The system requires and provides the highest levels of assurance of identity.

Principles 14, 15 and 16 – which are not represented in Table 11 – are added as a result of the risk analysis work presented in Section 7of this document.

# 10 EXEMPLARY MODELS

## 10.1 Overcoming privacy trade-offs

As we concluded in Chapter 7, implicit in some digital identity models is a trade off between privacy, on the one hand, and inclusion, on the other. Of all of the models considered in this report, national eID schemes and brokered identity systems currently strike the best balance between competing concerns. Although they are often plagued by more vulnerabilities than emerging and individualised identity systems such as Personal IDP systems, they are considerably more robust and secure than monolithic and federated identity systems. The risks they create are more easily mitigated through stronger regulatory and technical mechanisms. Over time it is possible that more distributed and user-centric approaches to digital identity will become mainstream but as yet they do not have widescale relevance.

As we emphasised in section 6, the impact of a particular digital identity system will always go beyond technological and commercial choices and depend in addition on the regulatory environment and accompanying governance mechanisms. As such, where national eID schemes and brokered identity systems are implemented in the absence of accompanying robust regulation and oversight, they will create further risks to privacy.

Identifying good models and best practices, therefore, requires an analysis both of the threats and vulnerabilities of particular systems, and the environment and manner in which they are deployed.

Using this approach, we have identified three good models upon which future digital identity systems could be based: Austria's Citizen Card scheme, the United Kingdom's Verify system, and the Estonian eID system. We have chosen these models not because they are free of negative impacts or risks, but rather because they have a number of positive features which could be replicated, and their negative aspects can be remedied by alterations to the system. We include in this section two further models; the Peru eID and India's Aadhaar system. Both systems have some positive features, but moreover are generally held out within the identity discourse as being exemplary and replicable models. We therefore analyse them alongside Austria, Estonia and the UK, although with the caveat that although we have included Aadhaar in our analysis, it does not necessarily fit squarely within the definition of digital identity embraced in this report.

## 10.2 Austria

| Positive aspects | Negative aspects |
|---|---|
| ✓ Comprehensive data protection law<br>✓ Independent data protection authority<br>✓ Limited data kept on the card<br>✓ Separation of identities by sector<br>✓ Integrates with 12 services | × Some concerns about the security of card readers<br>× No overarching body responsible for management |

### 10.2.1 Regulatory features

Because of the country's strong attachment to privacy, Austria's Citizen Card scheme is designed to provide a secure and privacy-friendly form of identity management. The Citizen Card is supported by a legal framework which includes the E-Government Act (which contains extensive regulation of identification and authentication in electronic communications with public bodies), e-government sector regulation, the Electronic Signature Law, regulation of the use of the Central Resident Register identifier, and the Data Protection Act. The legal framework is robust and enshrines pillars of the scheme such as the requirement that the use of a particular derived identifier for identification purposes shall be limited to that sector of State activity which is served by the data (s9 E-Government Act).

The strong legal framework is bolstered by the supervision and enforcement by a number of regulators, namely:

- The Austrian Data Protection Authority, which operates the register of PINs
- The Austrian GovCERT, operated by the department for Federal ICT Strategy of the Federal Chancellery, which manages security breaches and cases of e-ID theft;
- The Telekom Control Commission, which supervises the qualified signature framework.

One weakness of the system from a regulatory perspective is that there is no national authority which has the competence to give binding instructions to all of the involved institutions, including public administration as well as private sector participants, on all issues related to the Citizen Card (although the data protection authority has jurisdiction with respect to the privacy aspects). This may impede individuals' ability to audit the system and obtain redress when needed.

### 10.2.2 Technological features

Rather than comprising a specific token, the Citizen Card is a technology-neutral framework which defines minimum requirements that a token needs to fulfil to provide a signature-creation device that enables i) qualified electronic signature; ii) sector-specific identification, and iii) representation (i.e. the holder than carry out legal transactions on another person's behalf). Citizen Card tokens are available as health insurance cards, civil servant service cards, mobile phones, student cards and bank cards.[191] In general, electronic identity tokens in Austria are prepared for being activated as Citizen Cards, but the Citizen Card functionality has to be voluntarily activated by the holder and their actual identity data written on the token. The token contains minimal data about the individual: a qualified certificate, the name and date of birth of the individual, and a PIN derived from the Central Resident Register.

The critical technological feature of the Citizen Card which makes it a good model for emulation is the use of unlinkable sector specific identifiers (and associated cryptographic keys and digital certificates). For the same individual, each service provider uses a different identifier cryptographically derived from the Central Resident Registry number. This not only prevents the matching of individuals across their use of services, but enables the simple revocation and replacement of encrypted identifiers in case of fraud. It also minimises the amount of data held

[191] OECD, *Digital Identity Management: Enabling Innovation and Trust in the Internet Economy,* 2011, available at http://www.oecd.org/sti/ieconomy/49338380.pdf.

centrally, as each service provider retains their own identity data on the relevant individual, and is required to establish that independently of the data held on the Citizen Card.

### 10.2.3 Commercial features

Like the UK system, the Austrian scheme is designed to be taken up by both private and public sector service providers. The Austrian framework encourages public and private sector organisations to provide Citizen Cards to individuals and provides open source modules to online service providers to facilitate the integration of the card to their online applications.

One aspect which may impede the long term commercial viability of the Citizen Card is that it is not a mandatory requirement that individuals obtain one, which may lead to a low volume of acquisitions and thus lower the incentives for private sector involvement.

## 10.3 United Kingdom

| Positive aspects | Negative aspects |
|---|---|
| ✓ Comprehensive data protection law<br>✓ Independent data protection authority<br>✓ Decoupling of IDPs and service providers minimizes sharing of data<br>✓ Focus on end-user experience facilitates inclusion<br>✓ Innovation in privacy encouraged<br>✓ Contractual controls over identity providers | × Potential for hub to track usage and link identities<br>×Inclusion of "matching data set" in all identity transactions increases potential for tracking and surveillance |

### 10.3.1 Regulatory features

Over the past fifteen years, the United Kingdom has developed a strong regulatory environment to govern the acquisition and processing of personal information; the Data Protection Act 1998 follows EU standards and compliance with it is high, in part due to a robust and independent data protection authority, the Information Commissioner's Office, which takes a proactive approach to encouraging compliance with the Act, including by regularly issuing guidance on specific data protection laws. The independence of the Information Commissioner's Office (ICO) was most recently and vividly illustrated by the Commissioner's appearance before the Joint Committee on the Draft Investigatory Powers Bill in January 2016, where the Commissioner strongly expressed dissatisfaction with the government's proposed new surveillance powers due to their likely impact on privacy.[192] The ICO receives some government funding, but also receives fees from data controllers subject to data protection obligations under the Act, ensuring it is not entirely dependent on State funding. The ICO is able to issue monetary penalties to non-compliant data controllers up to £500,000.

[192] Danny Palmer, "Government shouldn't have 'willy-nilly' right to access citizen's private data, says information commissioner," *Computing,* 7 January 2016, available at http://www.computing.co.uk/ctg/news/2440921/government-shouldnt-have-willy-nilly-right-to-access-citizens-private-data-says-information-commissioner.

Verify is designed and operated to enable citizens to create digital identities with multiple IDPs. This is a positive aspect from the perspective of privacy, as individuals can keep a separation between interactions with different Government departments.

### 10.3.2 Technological features

The use of brokered identity providers, separate from service providers, is essential for reinforcing unlinkability and thus minimizing the ability for tracking and surveillance. While there have been concerns articulated about the use of the hub in the Verify system and its potential to enable tracking, these concerns could be mitigated through technological adjustments,[193] rather than simply having to place trust in the government's assertions that the hub would not be used to such ends.

### 10.3.3 Commercial features

Verify is based on the "service provider pays" commercial model, which is relatively positive from a privacy standpoint, as it minimizes the incentives for IDPs to analyse data for marketing purposes.

The ultimate objective is to create a marketplace for digital identity, of which the government is only one – and the first – customer. Should this approach succeed, it will ensure the commercial viability of the scheme long term. However, it may also have negative impacts on individuals' privacy, as it raises the possibility that personal information will be shared with other commercial customers of the digital identity scheme. The UK Verify programme has just commenced its consultation with the private sector,[194] so it is unclear yet how this aspect of the system will work. In any event, strong regulation will be necessary to prevent the onward sharing of personal information across multiple customers of the scheme.

## 10.4 Estonia

| Positive aspects | Negative aspects |
|---|---|
| ✓ Comprehensive data protection law<br>✓ Independent data protection authority<br>✓ Logs enable auditing<br>✓ Minimal data provided to service providers | × Some excessive data held on card |

### 10.4.1 Regulatory features

Estonia benefits from a strong European data protection law and independent regulator (the Estonian Data Protection Inspection) to ensure compliance. In addition, the Identity Documents Act was passed in 1999 to specifically regulate the issuance of the identity card, along with a separate piece of legislation, the Digital Signatures Act (DSA), which regulates the framework

[193] Luis T. A. N. Brandao, Nicholas Christin, George Danezis, and Anonymous, "Toward Mending Two Nations-Scale Brokered Identification Systems" (2015) 2 *Proceedings on Privacy Enhancing Technologies,* 1-22, available at http://www0.cs.ucl.ac.uk/staff/G.Danezis/papers/popets15-brokid.pdf

[194] https://identityassurance.blog.gov.uk/2015/09/30/private-sector-needs-for-identity-assurance-workshop-dates/

and rules required to effectively govern a national PKI[195] and digital signature infrastructure. The DSA also establishes requirements with regard to Certificate Service Providers, mandating a number of stringent requirements (financial and procedural), and Time Stamp Providers.

### 10.4.2 Technological features

Similarly to the Austrian system, in addition to advanced smart card features, the Estonian e-ID card includes an embedded PKI application which enables online authentication and digital signature with electronic certificates. Over a decade no security breaches have been reported.[196] More than 600 online government services are available through use of the online authentication system; companies have access to more than 2,400 services.[197] Services such as transport ticketing and voting are all available via use of the digital features associated with the eID. Minimal authentication data is shared by the ID provider with the service provider.

Interoperability in the system is met through the deployment of a common document format applicable to each service independent of its provider (DigiDoc) and a central common public resource, connecting national databases (X-Road).[198]

Although the data held on the card is minimal compared to some other systems, the card does display an individual's nationality and place of birth, arguably excessive information.

The certificates on the card are available publicly in a directory service and contain only the card holder's name and personal ID code, which are considered public data by law in Estonia. In addition, e-mail addresses in authentication certificates are also available in the directory. The directory contains only valid (active) certificates: if a person suspends or revokes his/her certificate, it is also removed from the directory and the data are no longer available. This public data source may introduce some privacy concerns especially as the data present provides greater opportunity for surveillance or tracking or makes citizen eIDs easier to steal (although this can be mitigated by making good use of the authentication capabilities that the citizen's smart card provides)

One positive privacy feature which enables greater auditing and oversight of the system is the retention of logs of all requests for authentication, which logs are available for the individual to view and scrutinize.[199]

### 10.4.3 Commercial features

The Estonian eID system is run through a closed public and private partnership. Two key private organizations work with the government to support the ID card project: Sertifitseerimiskeskus (SK) – a joint venture formed in 2001 between two of Estonia's largest banks (Hansapank, Eesti Ühispank) and telecommunications organizations (Eesti Telefon and EMT), which acts as the certificate authority; and TRÜB Baltic AS, now owned by Gemalto,

[195] Public Key Infrastructure; the management of cryptographic keys and certificates that underpins internet security and digital identity
[196] "Estonia Takes The Plunge," *The Economist,* 28 June 2014, available at http://www.economist.com/news/international/21605923-national-identity-scheme-goes-global-estonia-takes-plunge.
[197] "Estonia Takes The Plunge," *The Economist,* 28 June 2014, available at http://www.economist.com/news/international/21605923-national-identity-scheme-goes-global-estonia-takes-plunge.
[198] "Good Practice Case: eID in Estonia", *EGovernment – Interoperability at Local and Regional Level,* DG Information Society and Media, European Commission, October 2016, available at http://www.ifib.de/publikationsdateien/Interoperability_in_eID_in_Estonia.pdf.
[199] "Good Practice Case: eID in Estonia", *EGovernment – Interoperability at Local and Regional Level,* DG Information Society and Media, European Commission, October 2016, available at http://www.ifib.de/publikationsdateien/Interoperability_in_eID_in_Estonia.pdf.

which makes and personalizes the card itself — both physically and electronically. TRÜB receives the card application from CMB and manufactures the card, printing and engraving the personal data on the card, generating keys on the chip and embedding the certificates on the card.[200]

There are now more than 1,200,000 eID cards in Estonia. Citizens are required to purchase the eID card.[201] In addition, the government has introduced the concept of e-Residency, allowing any citizen from any country the opportunity to apply for an eID card which will become a digital identity they are able to use in other online contexts. This ensures that there will be continued commercial viability for the scheme.

## 10.5 Peru

| Positive aspects | Negative aspects |
|---|---|
| ✓ Comprehensive data protection law<br>✓ Independent data protection authority<br>✓ Acceptable level of technological security<br>✓ ID authority independent of the government | × No particularly strong privacy features<br>× No use of sector specific identities<br>× Cost may exclude poorer citizens<br>× Hampered by limited birth registration and missing birth certificates |

### 10.5.1 Regulatory features

Peru has a comprehensive data protection law that follows the OECD and EU standards closely,[202] established by Decree in July 2011,[203] which came into force with the issuing of Regulations under the law in May 2013. Interestingly, while the law mimics the European approach, according to trade documents the enactment of data protection law in Peru is attributable to commitments made under the US-Peru Trade Promotion Agreement.[204]

The law establishes the National Authority for Personal Data Protection, which may enforce compliance with the law and impose fines up to USD $135,700. However, there are concerns about the robustness of enforcement mechanisms and the strength of the data protection authority. As at July 2014, a year after the law came into effect, only 90 Personal Data Filing Systems had been registered in the whole of the country.[205] On 30 October 2014, the authority

[200] "Good Practice Case: eID in Estonia", *EGovernment – Interoperability at Local and Regional Level,* DG Information Society and Media, European Commission, October 2016, available at http://www.ifib.de/publikationsdateien/Interoperability_in_eID_in_Estonia.pdf.
[201] https://www.politsei.ee/en/teenused/isikut-toendavad-dokumendid/id-kaart/taiskasvanule/oluline-info-taiskasvanule-id-kaardi-taotlejale.dot#riigiloiv and https://www.politsei.ee/en/teenused/isikut-toendavad-dokumendid/digi-id/oluline-info-digitaalse-isikutunnistuse-taotlejale.dot#riigiloiv
[202] Graham Greenleaf, "The Influence of European data privacy standards outside Europe: Implications for globalisation of Convention 108, *International Data Privacy Law*, Volume 2, Issue 1, 2012 available at http://www.coe.int/t/dghl/standardsetting/dataprotection/Global_standard/GG_European_standards2010.pdf
[203] Data Protection law (Ley N° 29733)
[204] Available at http://www.sice.oas.org/ctyindex/USA/ftbper2009_e.pdf.
[205] Eliana Lesem, "Peru: Recent Updates to Peruvian Data Protection and Privacy Law," *Mondaq,* 28 July 2014, available at http://www.mondaq.com/x/330672/Data+Protection+Privacy/RECENT+UPDATES+TO+PERUVIAN+DATA+PROTECTION+PRIVACY+LAW

issued its first fine, of USD $78,000, against datos.org, yet reportedly was unable to identify the natural or legal person against which it could enforce the fine.[206]

Peru is a good example of a country in which a robust law, administrative independence and an independent data protection authority alone cannot be sufficient to ensure privacy law is complied with; sufficient financial and human resources and strong legal institutions must also reinforce the law.

### 10.5.2 Technological features

The Peruvian eID system is often held out as an exemplary model, particularly in the region. This may be in part due to the fact that in 2015, the Latin American Conference on Security awarded Peru the best ID card award at a ceremony held in Lima.[207] However, on our analysis (which is limited somewhat due to the lack of publicly available information about the technological architecture or implementation of the system) the Peruvian eID system does not have any specific technological features which make it remarkable or a particularly good model. While the scheme seems to deploy industry standard, sensible smart card technology, there are no innovative features which make that technology particularly privacy-friendly. There is no use of sector-specific identities, for example. The limited information on the technological model suggests that excessive information can be accessed through the eID, including information on marriages and divorces, for example.

The only positive feature in this regard might be said to be the independence of the eID authority, Reniec, which operates independently from the government and receives the bulk of its funding from outside sources beyond government.

### 10.5.3 Commercial features

The Peruvian system is government-led, requiring citizens to pay for eID cards. Because the citizen is the customer, it is expected that such a model is more likely to be consistent with protecting individuals' privacy.

One concern about the use of this model in a country such as Peru is the potential impact to inclusion of requiring citizens, particularly poor citizens, to pay for eIDs.

[206] "Ibero-American Data Protection Network Calls for Stronger Regulation; Peru DPA issues Fine," IAPP, 21 November 2014, available at https://iapp.org/news/a/ibero-american-data-protection-network-calls-for-stronger-regulation-peru-dpa-issues-fine

[207] Daniel Battu, "Peru's electronic ID card," *Smart Webzine,* 28 June 2015, available at http://www.smart-webzine.com/en/carte-didentite-electronique-du-perou-5351

## 10.6 India

| Positive aspects | Negative aspects |
|---|---|
| ✓ Major progress in rolling out digital identity service (Aadhaar) to over one billion Indians – an impressive achievement<br>✓ Strong technological security<br>✓ Focus on identity, rather than citizenship, facilitates registration process | × No comprehensive data protection law<br>× No independent data protection authority |

In recent years, India has made remarkable progress with the Aadhaar program in registering its people and issuing them with a digital identity, with many of these Indians being previously unknown to the Government. The program is ongoing; registration is continuing, and efforts are being made to 'seed' Aadhaar numbers into the databases of various social benefit schemes[208].

### 10.6.1 Regulatory features

The Aadhaar program was instituted on the basis of an executive order of 28 January 2009. However, it took until 26th March 2016 for the Lok Sabha (the lower house of the Parliament of India) to adopt the Aadhaar (Targeted Delivery of Financial and other subsidies, benefits and services) Bill 2016, which gives statutory backing to the use of Aadhaar for delivery of social benefits. It was passed as a "money bill", which means that the Rajya Sabha (the upper house) cannot block or amend it, the justification being that it deals with the way public money or other social benefits will be distributed under a range of government schemes. This Bill gives the Aadhaar program the legislative and regulatory grounding it was previously lacking for its use in e-Government.

Against the background of a very successful Aadhaar program, it is important to remember that India has neither a comprehensive data protection law nor an independent data protection authority, and negotiations about a privacy law have been under discussion for many years without significant progress.

The Bill includes a provision that the core biometric information (fingerprints and iris scan) cannot be shared with anyone for any reason whatsoever. There are no exceptions to this rule. Consequently:

- The Unique Identification Authority of India (UIDAI), the body responsible for the operation of Aadhaar, has been to court to challenge requests to use the biometric element of its register to support criminal investigations[209] on the basis that the UIDAI system has been developed "for civilian use and for non-forensic purposes".

[208] http://economictimes.indiatimes.com/news/economy/policy/indias-first-national-social-security-platform-to-be-developed-by-deity/articleshow/52301300.cms?from=mdr
[209] http://indianexpress.com/article/cities/delhi/stop-aadhaar-data-use-to-probe-crime-uidai-to-sc-2/

- Operationally, only online authentication is possible, and the matching is done at the Aadhaar servers – however, other non-biometric data held by Aadhaar may be shared with a service provider, on successful authentication.

  The requirement for online authentication is itself somewhat difficult, since there are many areas of rural India with either no mobile data coverage, or unreliable/insufficient capacity. Although an 'online only' policy is a laudable aim, it is currently impractical in some circumstances, and intermediate approaches using smartcards and offline biometric authentication will be necessary for some time.

Under the provisions of the Bill:

- The Aadhaar number can be used to assert the identity of a person receiving a subsidy or a service. If a person does not have an Aadhaar number, the Government will ask them to apply for it. Otherwise, the person will be allowed to use an alternative means of identification - once.

- Any public or private entity can accept the Aadhaar number as a proof of identity, for any purpose.

### 10.6.2 Technological features

During enrolment, an individual submits a series of biometrics (fingerprint and iris) as well as any supporting documentation about their biographical details (name, date of birth and address - information relating to race, religion, caste, tribe, ethnicity, etc. cannot be collected), known as "breeder documents". Where a prospective registrant does not have adequate supporting documentation, they can still be enrolled through the "introducer system" (two existing Aadhaar holders vouching for the identity of the introduced person who does not have the requisite breeder documents).

Based on the enrolled iris biometric, UIDAI undertakes various uniqueness checks to ensure that the submitted biometric information has not been entered into the Central Identities Data Repository (CIDR) previously, a process known as 'de-duplication'. Assuming that the presented identity is unique in the system, a new Aadhaar number is generated and associated with the submitted biometrics. The Aadhaar database minimises the amount of data held on an individual to fields including name, date of birth / age, gender, address and (optional) contact details (email / mobile), and specifically excludes details of race, religion, caste, tribe and ethnicity. It stores biometric templates for face, 10 fingerprints and iris scans, though only fingerprints are used for subsequent authentication.

There is a disparity between the numbers that are claimed to have been registered for Aadhaar, and the numbers in possession of an Aadhaar number, though no definitive figures are available. This disparity has arisen because of process: when a person is registered, they are issued with a temporary registration number; once the data has been returned to UIDAI, de-duplicated and otherwise validated, a new, unique Aadhaar number is generated, and printed on a card. This card is then delivered to the registrant. Since any postal service will on occasion lose packages, it is inevitable that some of the cards will go missing. The process for dealing with this situation is not well publicized, and many people in this situation have little or no idea of what to do about it, and don't know their Aadhaar number in order to initiate a query.

By focussing on physical characteristics such as fingerprints, UIDAI claims that the Aadhaar number is universal and always linked to the person as there is no physical card that might be lost or would need to be replaced.

Authentication of identity claims takes the form of various levels of Yes/No questions including: Is this Aadhaar number a validly issued number? Does this Aadhaar number share these biographical details such as name and address? Is this Aadhaar number associated with these fingerprint biometrics?

By linking the Aadhaar number to biometrics, the scheme minimises the risk of counterfeit fraud. If a fingerprint biometric check is undertaken, it is possible to ensure that subsidised food staples are only issued to those entitled to them; recording the Aadhaar number of those who have used their entitlement to 100 days of work can help ensure that these individuals cannot claim more than their entitlement. Similarly, the enrolment checks undertaken before an Aadhaar number is issued mean that it can be used as a suitable alternative to the proof–of–identity and proof–of–address documentation required to purchase telecommunications services (Department of Telecommunications, 2011).

Once an Aadhaar number has been issued, Aadhaar enabled bank accounts[210] can be opened if requested and the Aadhaar Payments Bridge facilitates seamless transfer of all welfare scheme payments to beneficiary residents[211]. Aadhaar numbers are also intended to be used with local "microATMs"[212].

Aadhaar eKYC has been accepted by the following regulators: Reserve Bank of India, Insurance Regulatory and Development Authority, Department of Revenue, Securities and Exchange Board of India and Pension Funds Regulation and Development Authority[213]

### 10.6.3 Seeding

If the Aadhaar service is to achieve its potential, then there is a clear need to use it to authenticate citizens when accessing services delivered across Government; this is an active consideration in, for example, the RSBY health insurance scheme, and the PDS food security scheme, and indeed in any e-government initiative. So there is a need to 'seed' the Aadhaar number into a range of existing databases, at the core of which are the National Population register (NPR) and the Socio-Economic Caste Census (SECC) (the ultimate aim being to derive the scheme-specific databases from the core Government identity databases, including Aadhaar numbers).

The mechanism for achieving seeding is as yet unclear, but until this is achieved the value of Aadhaar will necessarily be limited. At least two mechanisms have been discussed:

- Capturing Aadhaar numbers during the annual (re-)enrolment for social benefit schemes, and seeding to the scheme database and from there to the NPR and SECC. This method is likely to deliver accurate results, as it is automated, and would include an Aadhaar verification step to ensure that the correct Aadhaar number is being used, and that it relates to the individual enrolling.

[210] http://uidai.gov.in/fi-e-kyc.html
[211] https://resident.uidai.net.in/en_GB/aadhaar-services
[212] http://uidai.gov.in/images/commdoc/uidai_scheme_deployment_of_microatms_261012.pdf
[213] http://uidai.gov.in/fi-e-kyc.html

- An alternative being proposed relies on India's network of Accredited Social Health Activists (ASHAs), community health workers who are based in villages and on whom the burden of many rural initiatives falls. It has been suggested that the ASHAs could manually collect Aadhaar numbers from households, and upload them into the NPR. This would seem less than ideal, as it is prone to error through transcription and does not involve any form of verification that the Aadhaar number is truly that of the person whose record is being updated.

### 10.6.4 Commercial features

There are plans and pilots to roll out a range of different functions based on Aadhaar[214]. These include eSign that is intended to eliminate the need for "wet ink" signatures[215]. It is based on an open API that allows applications to integrate easily. In this way, an online request for a digital signature for an application would result in an eKYC check against the Aadhaar database and proof of address / identity verification from Aadhaar, such that the Aadhaar e-KYC service should provide digitally signed information that contains name, address, email address (optional), mobile phone number (optional), photo and response code to the applicant and the same should be shared with ESPs with the consent of applicant. Authentication to use this service would either be via biometric or one–time–password[216]

Another application area is a digital locker service[217], which allows digital documents to be stored securely, and access to the document by a requestor is controlled by the user. As Aadhaar is not compulsory, there are two ways to sign up for a DigiLocker account: "1. Aadhaar based method: You can voluntarily use Aadhaar (issued by UIDAI) to sign up using mobile OTP or biometric fingerprint device. And 2. Non Aadhaar method: You can authenticate your mobile number and then submit your proof of address and identity documents for manual verification". According to the website, there are just over 1 million users of the service by early January 2016 with 5 issuer organisations enrolled. There was a surge in interest in the service between May and July 2015[218].

UIDAI also intends to offer a unified payment interface using Aadhaar although the work of the payments council of India appears to be at the stage of API and technology specification[219].

More generally, these activities appear as part of a broader technological initiative known as the India stack, that intends to combine service layers based on consent, cashless and paperless transactions based on the "presence–less layer" that Aadhaar provides. This is also known as JAM – Jan Dhan ("people's money scheme"), Aadhaar and Mobile. In 2014, the Indian Government opted for the RuPay debit card over the Aadhaar-based platform for "last-mile" authentication in its Jan Dhan (PMJDY) service, due to the inability to use Aadhaar for offline beneficiary authentication (though beneficiary accounts are seeded with Aadhaar during enrolment). According to Microsave[220]:

[214] Presentation by Nandan Nilekani to Omidyar workshop, Brookings Institution, 16 November 2015
[215] http://cca.gov.in/cca/?q=esign.html
[216] http://cca.gov.in/cca/sites/default/files/files/e-AuthenticationGuidelines.pdf
[217] https://digilocker.gov.in/
[218] https://digilocker.gov.in/public/dashboard
[219] http://npci.org.in/documents/Int_API.pdf (February 2015).
[220] http://www.business-standard.com/article/current-affairs/last-mile-connectivity-a-bane-for-jan-dhan-scheme-116040200832_1.html

> *A third of the accounts opened under the PMJDY were not the beneficiaries' first accounts. Only 47 per cent of the people with PMJDY accounts have received their RuPay debit cards, while these have been issued to 85 per cent of the beneficiaries.*

This is an example of the distribution problem faced by a country as big as India; registering people, then creating and distributing the cards separately, at a later date, a task which is made even more complex by the need to distribute PINs separately from cards. The risk is that cards will then go uncollected in banks, due to the difficulty of travelling to collect them. And because PINs are sent separately, people need to visit a third time to collect their PIN envelope; so many go uncollected. And as a final blow there's a 3-month expiry on the PINs, so by the time a customer has his card and his PIN (after 3 separate visits to the bank), the PIN doesn't work anymore and the card is useless. It is therefore not a surprise that the success rate is as low as the reported 47%.

Finally, although we understand that UIDAI are to begin charging for Aadhaar authentications, there is no clarity either on when charging might begin, what level of charges are being considered, or whether it would apply equally to private sector organisations and Government programmes (though it seems likely that the latter will be exempt).

# 11 WIDER ISSUES

In this section we highlight some of the wider issues that need to be taken into account when considering digital identity services, including consideration of their impacts on privacy where appropriate.

A number of the issues highlighted in this section have considerably more relevance to emerging economies, as they relate to the maturity and reach of a country's infrastructure and the consequent effect on the establishment and operation of sophisticated identity services. However, many of these issues have resonance across all countries, whatever their stage of development.

## 11.1 Operability

There are some significant practical issues that affect the straightforward establishment and operation of a digital identity service. Consideration of these issues is broken down into the following structure:

11.1.1. Quality and coverage of data communications

11.1.1.1 Effects on scheme registration

11.1.1.2 Effects on subsequent authentication activity

11.1.2. Availability of suitable data centres

### 11.1.1 Quality and coverage of data communications

Data communications services, whilst improving, continue to present challenges particularly in many emerging economies, with coverage – particularly in rural areas – often being poor. Even in urban areas with good coverage, insufficient investment in telecommunications infrastructure often results in a capacity problem, so that a data connection may appear straightforward, but the connection is unreliable or has such a low data rate that it is unusable. These problems affect both registration and subsequent authentication for digital identity services.

Note that this section of the document is focussed only on the implications of data communications difficulties on the processes of registration and authentication. Questions around for example the desirability of mandatory registration, or the need to delay biometric registration for infants, are dealt with elsewhere in this document.

#### *11.1.1.1 Registration*

For registration, online access to digital identity databases is probably unavoidable, since there is a need to avoid multiple/duplicate registrations. To some degree this can be obviated by for example the offline capture of registration application details, which are subsequently uploaded to the identity service for cross checking and approval. However, this then separates digital identity token issuance – which would have to come later – from registration, which may increase costs and cause additional logistical problems, if the registrant lives in a remote, difficult to reach area and is particularly problematic for digital tokens.

There is then the question of biometrics. Ideally biometric profiles should be collected at the time of (offline) registration, but this can give rise to security concerns, so that biometric data is only collected at the time of token issuance, when it is added to the card and to the identity database. This concern is largely spurious – since the same concern should apply equally to the other data collected – but does remain a feature of heightened sensitivity around biometric data, particularly as biometrics cannot be revoked (although different templates could be generated).

In all cases, the offline collection of registration data, potentially including biometric data, gives rise to privacy concerns, since the data is necessarily being held on a portable device carried by an appointed registration officer. It is therefore important that the registration data, once collected, should be held on the device ONLY in encrypted form, such that the data is not available to either the holder of the device or anyone else, and can only be decrypted when it reaches the data centre. Care must also be taken to ensure that temptations to collect additional data, unrelated to the registration process, such as location data are not satisfied through careful design of the devices and their systems.

Consult Hyperion has piloted such an approach to identity management in Nigeria, as part of our TAP Programme, as described in Section 11.8, Appropriateness. It was planned after that pilot that TAP be extended to support pre-registration for the Nigerian NIMC service, with distribution of the NIMC card itself remaining the responsibility of NIMC. However these plans were curtailed when the Nigerian Presidential elections intervened.

#### *11.1.1.2 Authentication*

Communications limitations are not an issue where a digital identity service includes the ability to do offline citizen authentication. However some services, such as Aadhaar, require online authentication, and this causes problems for residents in a range of circumstances. For example, it is known that the operators of a number of social benefit services across India would like to ensure they are offering services only to properly enrolled citizens, but cannot use Aadhaar because the service offers only online authentication; they are therefore considering using the foundational Aadhaar identity service as the basis of their own functional or transactional digital identity service, which is able to support offline citizen authentication through the use of a smartcard. However, this would require that they carry out their own biometric registration process, as it is not permissible to use Aadhaar biometric data outside the Aadhaar service. Alternatives to biometrics, such as PINs, are not acceptable, as it is the received wisdom that beneficiaries of social programs cannot use PINs, which experience has shown is not necessarily the case.

This of course gives rise to concerns around privacy as well as the cost of systems that duplicate Aadhaar functionality, and any such functional or transactional identity service must take care to address concerns in this area.

### 11.1.2 Data centres

It is unfortunately the case that a significant proportion of emerging economies – though certainly not all – do not have data centres that are suitable for use in support of digital identity services, still less the two that are necessary for robustness and disaster recovery (one data centre being designated ‘live’, and the other a ‘standby’, able to takeover seamlessly if the live data centre encounters a problem). It is commonplace for many countries to have issues around power, which are addressable through the use of generators (though this is an expensive

option, driving electricity costs as high as 10 times that typically available in the US[221]); of greater concern are issues around connectivity (addressed above) and security.

It is unfortunately the case that many countries do not have an effective, reliable approach to physical security; for example, we are aware of frequent cases in which data centres, where they exist, are easy to access through doors propped open with piles of paper, fire extinguishers etc., for reasons of convenience or even simple ventilation; where front desk staff are so welcoming that they do not think to ask for credentials; where metal detectors are used as a piece of theatre, whose only purpose appears to be to make the machine beep. A significant cultural attitude change is needed in many cases. Important-looking people with briefcases are not always benign in their intent.

Certification is not necessarily a guarantee of quality. An example is the Abuja data centre of the Nigerian national identity service, which is certified by a major international payment scheme, but the physical security of which would not meet the demands of that international payment scheme if it were sited in Europe or North America.

All of these issues give rise to privacy concerns. At the most basic level, it is not inconceivable that in some cases a straightforward social engineering approach would be enough to gain access to a national identity database, allowing copying, amendment or deletion of either single records or indeed en masse.

The above issues can be addressed using a cloud-based approach, or more likely (for enhanced security) through the use of a certified data centre elsewhere in the world with good connectivity into the country in question. However hosting a national identity scheme (or backup) outside the country it applies to probably contravenes local regulation and is invariably a matter of national pride and, depending on the location of the data centre adds additional risks (cf Schrems and Safe Harbor). There may also be cost and latency issues, particularly if facilities in nearby countries are not much better than in the host country.

## 11.2 Commercial Case

There is generally (initially at least) no commercial case for a digital identity scheme. Indeed, this is rarely the principal driver behind the establishment of such a scheme – drivers are more generally around governmental objectives, including the planning and delivery of government services and social benefits.

However, such a digital identity scheme offers significant value as the root of registration for other services such as travel documents (passports), health, education, training etc, through the development of a functional digital identity.

Such functional identities can be useful for public bodies, NGOs and private organisations involved in the delivery of services – for example, where profitability (or continued funding) is dependent on reporting of results, since Monitoring and Evaluation (M&E) capabilities can easily be built on top of the digital identity service. This adds weight to the reporting function of such services, and so validates their effectiveness – a valuable service.

[221] http://www.economist.com/news/middle-east-and-africa/21685504-electrification-plans-are-stalling-because-distributors-wont-pay-power-hungry

So as well as the traditional approach of charging for passports, civil registration certificates, etc., in the longer term there is potential for a strong commercial model that generates substantial income for the operator or owner of a functional digital identity scheme, through, for example, fees per authentication. This must be set against the privacy challenges this presents, though a properly designed service, which only divulges identity and personal data when necessary and relevant, is perfectly possible.

## 11.3 Liability

According to a recent White Paper[222]:

> *Key to expanding the online identity market is knocking down barriers to entry. Fear of liability is one of those barriers. But in this emerging ecosystem, fear of liability is little different than a fear of the unknown.*

In their paper, Smedinghoff et al, seek to establish a common understanding and vocabulary of basic liability considerations. They use the straightforward definition of liability as "the legal obligation to pay money to compensate someone else for the losses they have suffered", and make the following definitions:

- *The legal rules that determine when, and under what circumstances, one party is liable to another exist on three different levels: (1) general public law, (2) identity-specific public law, and (3) contract-based private law;*
- *At all three levels, those liability rules reference a "duty" or "obligation" that is imposed on a party. Thus, the first step in assessing the liability risk assumed by any business in an identity transaction is to identify and understand the legal duties imposed on it in connection with the role it undertakes in such transaction;*
- *At all three levels, those liability rules apportion liability to the person who is "at fault," by breaching a duty imposed on it, except in certain limited situations where the rules impose liability without fault (or strict liability); and*
- *In all cases, while there are differences in the way the damages are calculated, the basic goal is to compensate the injured party for the losses it has suffered.*

These definitions need to be considered in the light of the different models for digital identity discussed in this document. For example, the UK's Verify programme has a "buyer beware" liability model, which has clear advantages for the identity provider, but significant disadvantages for the less issue-aware citizen – so normal commercial liability arising from for example negligence applies, but that liability is not linked to the value of the transaction. By contrast, IdenTrust takes a more nuanced approach, with liability being a matter of negotiation and contractual agreement between a participant bank and a relying customer.

But in many emerging economies the regulatory underpinnings for digital identity schemes lag the capabilities of the technology. Further, even where the regulatory framework is in place, enforcement is often an issue, as described in Section 11.7.3 Enforcement, and this can render questions of liability irrelevant.

[222] "The Vocabulary of Identity Systems Liability", Thomas J. Smedinghoff, Mark Deem, and Sam Eckland, June 2013

## 11.4 Scale

### 11.4.1 General

Scale is a significant issue for digital identity schemes across the larger countries, including the emerging economies. Although this is not an issue for smaller countries, a country like India with a population of 1.25 billion (almost four times the size of the USA) faces significant challenges in designing, deploying and operating a digital identity scheme:

- The scheme has to be capable of handling the number of records involved and of rejecting duplicate registrations whilst accepting that in a large population many apparent duplications are in fact legitimate registrations thrown up by the workings of probability.
- The volumes of associated authentication transactions (for online authentications) will be similarly significant, and should not be underestimated. For example, if everyone in India authenticated themselves once a week on average, that would result in authentication requests (assuming a 12 hour day) being received at an average of over 4,000 per second; the peaks would certainly be even higher. These volumes are not insignificant, and require careful design of the service and its capacity.
- The simple task of reaching all of the citizens in a country with infrastructural challenges in remote rural areas (such as unmade roads that disappear in the rainy season) should not be underestimated – and these are the areas most likely to be out of the reach of communications networks.

To these must be added the challenge of multiple languages – India has 22 official languages, and there are said to be more than 1,650 languages in common usage. This causes significant problems, for example when using teams of registrars with experience of registering citizens in one region in another region where registrar and citizen do not share a language. This issue extends beyond the initial registration to support in subsequent usage, registration changes, investigation and dispute resolution, etc.

### 11.4.2 Contracting out

The sheer scale of large identity services often gives rise to the need to contract commercial organisations to carry out registrations, an approach which is potentially fraught with problems. There are obvious privacy concerns; further, a commercial organisation, paid by the number of registrations completed, is unlikely to focus on the quality of those registrations, and will instead focus on speed of registration, even if minimum acceptable standards are set before registrations are paid for. This is an issue, amongst other things, with biometric registration, since reducing the quality of biometric capture increases the ease of capture and therefore its speed, but the completed record is of lower quality and the biometric captured may be of limited value.

These issues can be addressed using technology, automating the process as much as possible, reducing the opportunities for cutting corners, and ensuring all data is encrypted and cannot be misused.

## 11.5 Inclusion

### 11.5.1 Political

Inclusion should be the aim for every identity scheme, digital or not; but as mentioned elsewhere in this report this is not always the case. Since once an identity scheme is in place it would rapidly become the basis of voting entitlement (supplanting long term, but largely unsatisfactory, approaches such as voters' cards), in the majority of cases giving someone an identity card explicitly gives them the right to vote.

So where there is insufficient political differentiation between a registration agency and a dominant political party, there can be reluctance to register and issue cards to political opponents, which can lead not only to disenfranchisement of significant proportions of the population, but also as an (unintended?) consequence lack of access to social benefits such as pensions. Similarly, trust in a card issued by the previous government might be undermined by the election of a new government. This also applies to any benefits available from international NGOs, who are likely to rely on the existence of an identity card.

### 11.5.2 Financial

Due to international pressures around countering the financing of terrorism[223], financial inclusion is itself increasingly contingent on the holding of some form of national identity card, digital or not. So political exclusion can lead directly to financial exclusion, since access to bank accounts and (derived from this) other financial products becomes contingent on the citizen having an identity card.

However, there are often complications around this, and commercial banks frequently feel themselves vulnerable to prosecution if they get their approach wrong; resulting for example in the situation in Nigeria, where the new NIMC card is not accepted for registration for a new bank account, because the national financial regulator has not added it to a list of accepted forms of identity documentation (this is not to be confused with the use of a National Identification Number, NIN, which is issued at time of registration in lieu of a card, and which IS accepted – indeed, it is mandated by NIMC).

### 11.5.3 Surrendering privacy for finance

There are a range of initiatives across emerging economies, where typically organisations such as credit reference agencies do not have a great deal of penetration (and certainly not amongst the underbanked population), for the development of alternative approaches to reducing the risk in lending and for developing products aimed at financial inclusion. These alternative credit reference services use approaches such as analysis of mobile phone records to determine airtime spend, which is assumed to be discretionary expenditure and therefore an indicator of funds available to repay a loan.

In more developed economies, we see an increased use of social graph data (see Section 5.1.3.1) to identify people, not only for registration for digital identity schemes, but also in applications for financial service products – in which case social media activity is additionally being used to derive an estimate of the creditworthiness of the applicant, in the context of 'likeliness to repay', quite distinct from the ability to repay, and is an adjunct to the traditional credit scoring report.

[223] See the FATF Recommendations http://www.fatf-gafi.org

The concern here is the potential misuse of this data (there are already concerns around its use to vet employment applications). But concerns go further when considering the underbanked, wherever they are in the world; someone in urgent need of finance for whatever reason – such as a family emergency – can be pressured into revealing private information to an organisation in return for access to funds.

## 11.6 Interoperability

For the majority of emerging economies, interoperability is initially not an issue, as it is viewed as an end in itself, and in any case there are few e-services (governmental or otherwise) that could make use of it. However once a scheme is in place, it is typically the case that others seek to use it, to develop a range of transactional or functional identity services that rely on the foundational identity scheme, and either extend its capabilities or customise it to meet the needs of a particular segment (such as the already-mentioned possible offline services which use Aadhaar during enrolment). It is therefore clear that a scheme which is not interoperable, and which forces each e-service to develop its own solution, will quickly become inadequate. For this reason, interoperability should be viewed as paramount from the outset.

There is also the broader need for interoperability with, for example, banking services that rely on the digital identity for registration for financial services and subsequent access. This applies across the spectrum of 'client' services. It is these cases that may give rise to privacy concerns, since the digital identity service must ensure that only necessary, appropriate information is provided, and this should be pseudonymous (or assert a characteristic, such as 'aged over 16') where this is appropriate. The use of open standards in this regard will assist in both the process and the addressing of privacy concerns.

## 11.7 Funding

Funding for a digital identity service or scheme is always an issue, not least due to the scale of the initial registration task and the on-going operational costs. This is especially the case in many emerging economies, where governmental budgets are often tightly constrained.

### 11.7.1 Development

It is commonplace that the governments of emerging economies are encouraged to undertake the development and subsequent rollout of identity cards and services. This direction can often be traced to the international FATF Recommendations, and the consequent need to identity each of the parties to a financial transaction. Since it is to some degree in the international interest, governments are often provided with some element of funding for the development of a national identity scheme by international agencies, by means of some form of grant, often backed by some form of practical support in developing the associated regulations.[224] This will generally result in a significant degree of success in the development and initial rollout of a national (digital) identity scheme.

[224] Such funding is also commonly available for the development of a functional identity registration, for events such as national elections. However, the basis of registrations is commonly less reliable than that required for a foundational identity scheme, and reliance on it for the derivation of further functional or transactional identity schemes cannot be considered best practice.

### 11.7.2 Operation

Once a digital identity scheme has been established, there are significant on-going costs in its operation and maintenance, both for the service itself (staffing and general operational costs) and the citizen records it holds (at an absolute minimum, the issuance of new cards as citizens reach the appropriate age, (where necessary) the updating of citizen data held centrally and the subset of that data held on any card according to life events, replacement of lost cards, and the cancellation of cards when a citizen dies).

The UK Verify service's approach to this is to target registration of 10% of citizens per year, with each registration having a 10 year validity – so after 10 years, the entire target population is registered, but those first registered must be 'refreshed'. By contrast, there is no evidence that the Indian Aadhaar service has capacity for such a 'refresh' operation, and concerns around funding are likely to be at the root of this omission.

However, ongoing funding is often assumed by international agencies to be the remit of the national government, and such funding is commonly inadequate. This can have significant consequences for the long-term viability of a digital identity scheme; provision of registration centres to support ongoing registration updates can be patchy, and any national identity database quickly becomes out of date and devalued. Data centres can become less robust as costs are cut, and any online services can suffer impacts on their availability, so damaging the reputation of the service. Security is often one of the first things to suffer, with obvious potential consequences for both privacy and the integrity of the national identity database.

### 11.7.3 Enforcement

As has been noted elsewhere in this document, no digital identity scheme has viability without appropriate enforcement of the associated regulations, particularly with regard to privacy.

As per the comments for on-going operational costs, funding for enforcement is generally the purview of the national Government, and tends to slide. It is unfortunately the case that international agencies are more commonly interested in funding the development of a new flagship project, rather than providing long term, multi-year funding for the operation of an established service, or the enforcement of the legislation around that service.

## 11.8 Appropriateness

The funding of a national digital identity scheme by an international agency has the potential to result in the development of a scheme which meets the needs of the international community, without detailed consideration of the needs of the country and its citizens. So in a country with few digital services – either local or remote – it is arguable that a full-blown digital identity service is not strictly necessary, and a conventional card-based service, with a roadmap for evolution to a digital identity service over a period of 10 years or more, might be more suitable for the country, considerably more affordable, and actually achievable in a country with challenging physical and networking infrastructure.

Set against this are the needs of the international community: a robust digital identity service, able to validate the identities of all parties to all financial transactions (including access to basic financial services); and the need to monitor the effectiveness of international support or other interventions, through an auditable M&E approach to projects.

It is of course in the long-term interests of emerging economies that they evolve to a fully connected digital economy, supported by a flexible digital identity service. Whether or not in every case they truly need a world leading digital identity service with an ongoing operational and enforcement cost they cannot really afford is a moot point – particularly since many relatively developed economies have managed perfectly well without one until their economy has developed sufficiently to need one.

An alternative approach might be similar to that undertaken by Consult Hyperion in Nigeria, where the TAP Programme[225] was used to register farmers in remote, rural Nigeria, well away from communications networks. Our team of registrars would collect identity and demographic data about the farmer using an Android tablet, and at the end of the process issue them with a contactless TAP card, which they could use to identify themselves to access further services. The farmer data collected would be transmitted to a central farmer database, and used to target social benefits based on data collected. Whilst not a foundational identity service, TAP represents a transactional identity service that is appropriate to the needs of the citizens, rather than being focused on the needs of international agencies.

#### *11.8.1.1 Multi-functionality*

Many of the more ambitious digital identity schemes and services offer significant degrees of multi-functionality (a prime example being the Nigerian NIMC service). It is reasonable to repeat many of the comments made above regarding appropriateness to this multi-functionality, the majority of functions of which will never be used by citizens.

[225] http://www.chyp.com/token-administration-platform-tap-e-goods-delivery/

# 12 BACK TO THE BIG PICTURE

By way of conclusion we present five questions to ask when considering privacy friendly digital identity systems:

## 12.1 What is the (digital) identity system trying to achieve?

Perhaps the most fundamental question to answer is what policy goals the identity system is trying to address[226]. Whilst it is recognised that additional use cases will emerge, particularly once an effective digital identity infrastructure is in place, a specification of the core requirements will shape the overall design of the system. For example, if one of the key goals of the identity system is to reduce fraud and costs of making secure electronic transactions, it makes sense for all *residents* in a country to be able to enrol into that system and such a system is most likely to be useful once individuals reach the age of majority. In contrast, a system that is focussed around the rights and obligations of *citizens* might be more closely linked to existing birth registration details. This question is particularly important in the context of moves towards fully digital identity systems that can have significant investment and infrastructure cost consequences, particularly for emerging economies.

### 12.1.1 To what extent would identity credentials address the stated policy objectives?

One frequently overlooked part of this question concerns the extent to which there is a proper understanding of the role that mis-identification plays in the policy goals being addressed. For example, benefit fraud and associated "leakages" are frequently cited as arising from mis-identification problems, e.g. an individual who has been (wrongly) issued with multiple ration cards. In many cases, however, a closer investigation reveals that the fraud has less to do with incorrect claims about identity and more to do with incorrect claims about circumstances (e.g. misreporting income levels). Even if the potential for fraud and corruption is directly related to identity claims, the wider context of use needs to be understood to determine the extent to which it is possible to undermine / ignore high quality identity credentials (i.e. failing to use effective authentication). For example, a context where entitlement to subsidised food is associated with a unique number and a verified fingerprint would only work if the *only* way for the food to be issued was once a verified fingerprint and entitlement was authenticated. Any opportunities for "workarounds", perhaps because "there are connection / power problems" would undermine the effectiveness and public trust in the system and enable "rent–seeking" behaviours.

### 12.1.2 How transitive is the trust in existing credentials?

In many situations, trust is assumed to be transitive: IF A trusts B and B trusts C then A should trust C. In practice, and particularly when commercial decisions are being made, the limits of this transitivity become apparent. One illustration of this can be seen in the relationship between functional and foundational identity systems. Functional systems, such as those to support democratic elections are often strongly supported by international development agencies and other funders. Entitlement to vote is an attribute closely linked with notions of citizenship. Logically, therefore, a well implemented voter registration system should provide a strong evidence base either for issuing foundational documents to citizens or, at very least,

[226] Whitley Edgar A. and Gus Hosein (2010) Global Identity Policies and Technology: Do we Understand the Question? Global Policy 1(2), 209-215. (ISSN 1758-5880)

providing a strong evidence point for identity verification activities. In practice, however, even in situations where there are not constitutional restrictions on reusing election data for other purposes, this kind of transitivity doesn't always apply.

The same issue arises in the use of credentials for commercial purposes. Does an appropriately issued, genuine identity card provide sufficient proof of identity to satisfy national and international KYC requirements? If there are queries about identity claims, where does the liability lie? With the identity provider (who may be a national government) or with the relying party (who is, by definition, unable to rely completely on the claims being supported by the identity credential)? This is discussed further below.

## 12.2 What levels of assurance are needed?

Not all identity related transactions require the same level of confidence in the claims being made by the individual and this distinction applies as much with digital identities as with analogue identity credentials. Claiming to be the same person as the one who started an (online) application process does not have the same level of risk as claiming to be the person entitled to direct cash transfers from government. These different kinds of transactions will require different levels of assurance to support the claims being made. Whilst it is possible to insist on the highest levels of assurance for all transactions, this can be exceedingly costly, require high levels of infrastructure roll out and introduce unnecessary privacy risks to the process. A more nuanced, risk based approach to the question has clear benefits for both developed and emerging markets and is likely to reduce the extent of attempts to work around apparently pointless and inconvenient identity activities.

Some identity systems are presented in terms of absolute certainties ("gold standards of identity"), not least because this has an intuitive appeal. Achieving gold standards of identity verification and authentication is likely to be very difficult to achieve, incredibly costly and unnecessary given the different levels of risk associated with particular transaction types.

Note, this is not to say that higher levels of assurance should not be supported by an identity system but rather that it may not be necessary to ensure that all members of society, from the poorest agricultural workers to senior politicians need to undergo the same identity verification processes. Similarly, once an identity has been verified, different kinds of transactions might require different levels of authentication. In this context, biometric uniqueness (as is the basis of the Aadhaar scheme) does not provide a full gold standard of identity, even if it does (within the tolerances of the chosen technology and its implementation) demonstrate uniqueness.

There is growing recognition of the need to standardise levels of assurance, including identity verification, to ensure their interoperability. In the UK, the Verify GPG 45 provides guidance on the requirements for identity proofing and verification of an individual using online services[227]. Similar standards are being developed for the US NSTIC programme and the EU eIDAS regulations.

### 12.2.1 What identity evidence is required for particular transactions?

Questioning the level of assurance needed for particular transactions raises the associated question of what identity evidence is required for a particular transaction. Here, the data protection principle of data minimisation is particularly effective for enhancing the privacy of

[227] https://www.gov.uk/government/publications/identity-proofing-and-verification-of-an-individual

most identity related transactions. In many situations, the answer to the identity evidence question is Yes or No, not "Full name, date and place of birth, gender, address" etc. In the case of analogue identity cards, it makes sense for some of this data to be presented in human readable form, despite the risks[228], however, digital identity systems can be designed to support privacy friendly disclosures.

Effectively, what this means is that for many identity related transactions it is not actually "identity" information that is required, rather it is identity-related attributes, particularly when combined with effective risk assessment. Thus, accessing age related products doesn't need identity information, it needs a Yes / No answer to the question of whether the individual in front of you satisfies the age-related criteria ("is the person over 18?", "is the person entitled to a pensioner's discount?")[229]. Although age related claims are easily understood, the same logic can apply to claims about nationality, entitlement to use social services, the right to vote etc. Another illustration might be the distinction between a "driving licence" (UK) and a "driver's licence" (US), even if in practice both end up being primarily treated as government issued identity credentials with photos.

Whilst age related entitlements are frequently open–ended (once you are 18 years old you will always be at least 18 years old), other entitlements / attributes are more dynamic. Thus, an individual may only be entitled to subsidised travel whilst they are unemployed. There are two ways of addressing this: First, by making a reasonable (risk based) assumption about the minimum period that this entitlement is likely to be apply. Thus, it may make sense to give someone a (renewable) six month entitlement to subsidised travel because they are receiving unemployment benefits at the time of issuance than to attempt to link up a real-time checking of ongoing entitlement for every journey. The potential cost of nominally ineligible subsidy being consumed is easily outweighed by the savings in real-time infrastructure checks. An example from the commercial sector can illustrate this: although payment systems include the functionality to check that every credit card transaction is being made by a valid card that is still within its limits, in many situations (e.g. busy supermarket checkouts), a commercial (risk based) assessment is used to "approve" low value transactions without doing this check. A similar situation applies with the £30 no–PIN limit on contactless cards in the UK.

Risk–based attribute checks also allow for the possibility of "zero knowledge" checks to be made. For appropriate transactions, where the relying party has confidence in the attribute claims being made (e.g. in a bar, an individual satisfying the claim that they are over 18) and with a suitable legal / regulatory environment, there is no need to keep an auditable check of the transaction to be kept on some database. There is increasing evidence that the meta-data stored in such audit databases can have serious privacy consequences as they allow significant inferences about a person to be made.

Paraphrasing examples about telephone metadata from the EFF[230], if there is an audit trail record of your identity being checked at a gynaecologist's office and, later that day, an audit trail record of your identity being checked at a local Planned Parenthood's office, strong assumptions about recent life events can be inferred.

[228] http://digitaldebateblogs.typepad.com/digital_identity/2009/10/what-a-cunning-stunt.html
[229] Birch DGW (2009) Psychic ID: A blueprint for a modern national identity scheme *Identity in the Information Society Open Access Journal* Archived at http://dx.doi.org/10.1007/s12394-009-0014-6
[230] https://www.eff.org/deeplinks/2013/06/why-metadata-matters

Zero knowledge proofs are not appropriate for all circumstances, in some cases there needs to be real-time checking of attribute claims ("is this person (still) an employee?") or where there is a clear need for audit trails of identity verification. These are, however, policy choices and it is better that these decisions are the result of explicit design decisions rather than being the (privacy unfriendly) defaults.

It is recognised that appropriate implementation of zero knowledge proofs is unlikely to be straightforward as there needs to be a wholesale change in risk understanding and attitudes as well as, possibly, a change in the legal environment regarding liability

Tokenisation / pseudonymisation provides an alternative solution where some linkability is required (for audit related activities) but the linkability, in normal circumstances, doesn't need to be back to an identifiable individual. For example, the shibboleth federated access system for universities allows students (and staff) to access online resources like electronic journals. Nominally, the publishers of these journals only require confirmation that the person is entitled to access the materials ("is a registered student or staff member") and this could be based on time limited claims (e.g. a single token that is used by all students for the whole academic year). In practice the claim is associated with a random token associated with the particular student. In most cases, the relying party is satisfied with the institutional support for the claim that the person is a student. However, in some circumstances, for example, if the student is posting this material freely online, the relying party can use the token to ask the university to "decrypt" the token and re-identify the errant student.

### 12.2.2 What is the best way to maintain the integrity of the identity credential?

If the identity credential has value in the society (e.g. for accessing public services), then there is inevitably a risk of the credential being compromised. This can take a variety of forms: "fake" credentials might be produced or "genuine" credentials produced but issued to, or used by, the wrong person. These problems can be associated with both analogue and digital identities, although the solutions are very different.

The problem of "fake" analogue credentials is increasingly being addressed by ensuring that they include features like Guilloche printing (a decorative engraving technique in which a very precise intricate repetitive pattern or design is mechanically engraved into an underlying material with fine detail), microtext, serial numbers printed on the face of the card, a read-only chip on the card that would also store the card's serial number or optically variable ink that changes colour depending on the viewing angle. Adding these security measures, however, also requires training staff who know how to check them (at least in terms of basic checks) and having appropriate back office capabilities for escalating queries if suspicions about the genuineness of the credentials are not resolved.

For digital identity credentials, "fake" credentials should not be an issue as cryptographic techniques will reveal, if checked, that the credential was not appropriately issued (i.e. not signed by the private key of the issuing body).

The issuing of "genuine" credentials to ineligible individuals is more problematic and requires a holistic view of the whole issuance process to address potential weaknesses in the process. One advantage of digital identity credentials is that if it is discovered that a particular credential (or set of credentials) is invalid, it may be possible to "switch them off" (e.g. by rejecting any claims made using them) if they are checked against a central register.

The final problem arises when genuine credentials that were issued correctly are used by someone other than the intended owner. For analogue credentials, this typically requires checking, for example, the 2D fingerprint barcode with the live fingerprint of the person presenting the credential. In many situations, this authentication process is not undertaken as the underlying transaction only requires a low level of assurance. Emerging digital technologies, however, allow for the potential of higher levels of authentication assurance to be easily made, even for lower level of assurance transactions. For example, the fingerprint readers on new mobile phones (which only provide a one–to–one match between the "owner" of the phone and the presented fingerprint) could provide high levels of confidence that the person presenting the identity information (perhaps an Aadhaar number) was the person to whom the information was associated. Note, this example does not involve checking the fingerprint against the Aadhaar registry (which introduces privacy risks in terms of the associated audit trail of verifications), rather it would require the mobile device and Aadhaar fingerprint to be securely bound once. Thereafter, use of the live fingerprint would provide "zero knowledge" proof that the Aadhaar number belonged to the person whose fingerprint unlocked the phone.

## 12.3 Why go digital?

At this point it is perhaps helpful to reflect on whether it is, in fact, worth going digital for identity, particularly in developing economies where the need for secure online access to government services is likely to be relatively low. As discussed in this report, digital identity systems provide effective assured mechanisms for identification, authorisation and authentication. These do, however, come at a cost in terms of implementation, particularly in terms of roll out of digital credentials and readers. For emerging economies, the level of assurance required for many (current) transactions may not therefore require the roll out of digital identities and spending money on them at this time in the belief of "leap frogging" analogue identity cards will solve key development challenges. A digitally enabled identity that is infrequently authenticated electronically (as is the case with large numbers of digital smart-card enabled identity cards in both developed countries and emerging markets) and is based on poor quality biographical data (and, possibly, uniqueness assured biometrically) might provide a "bright shiny" high profile project that is well received by national leaders, funders and the shareholders of identity system vendors but achieves little towards goals of inclusion and development. In many cases, there is still space for analogue, physical cards that can be checked by eye (possibly in conjunction with under-the-table cash) in circumstances where connectivity to central registers is unreliable. Alternatively, if there is a move towards digital, it is possible to roll out smart card readers that can perform local authentication checks against details held on the chip card without necessarily needing a connection to the central register. Unless integrated with other capabilities (e.g. payments) such readers might end up being just an additional piece of kit that needs to be procured (and refreshed periodically). However, there is growing take up of portable computational devices that include secure elements and can provide high quality authentication mechanisms, namely mobile phones. These devices increasingly can provide high quality authentication mechanisms (discussed in 5.1.3.2) and whilst the device might also contain other data that has privacy consequences (such as location) there are technical measures that can minimise the risk of this data being disclosed in the transaction.

### 12.3.1 What is the role of mobile?

As noted above, the most advanced mobile devices do include the capabilities for sophisticated authentication techniques, for example, based on biometrics. Earlier generations of phones,

with the secure data element of SIM cards and one-time-code SMS codes can also provide relatively high levels of digital authentication for many transactions.

It must be borne in mind, however, that although the total numbers of mobile devices is rapidly reaching the same numbers as the global adult population, their distribution is still distorted with inequalities in mobile phone access currently mirroring existing inequalities in society (including by gender, literacy and urban / rural living) although these demographics are shifting rapidly. At this time, therefore, it is probably unwise to make a strong link between a particular mobile device and a particular individual as many phones are still linked to families or communities rather than individuals. The cost of the device, of calls and the quality of coverage, particularly for data also affects the potential for mobile to transform this space.

A final consideration, of course, is that in many countries, some form of identity credential is needed before SIM cards can be issued to individuals.

### 12.3.2 What is the role of biometrics?

Biometrics (typically fingerprint, iris or face) can be used in two very different ways in digital (and analogue) identity systems. The first use involves one-to-many matching. Here a given biometric (more accurately, the template generated from the biometric) is checked against all existing biometrics held in the identity system to see whether that biometric (and, by implication, that person) has already been registered in the system, indicating that an attempt is being made to create two different identities in the system for the same person. Whilst uniqueness is frequently presented as a desirable policy goal (assuming that it is matched with appropriate authentication methods), it is not a neutral proposition. Any biometric has known performance rates and, inevitably, some individuals will not be able to provide biometrics of a suitable quality. Biometric template matching also raises practical concerns when near matches are discovered and need to be checked manually to determine whether they legitimately come from different people with similar biometrics or are part of an attempt at fraudulently registering two identities for the same person.

One-to-one biometric matching can be more flexible as it is only checking the (reasonable) assumption that the person presenting the biometric is the one who registered the biometric originally. Apple's TouchID (and the Android equivalents such as Nexus Imprint) are perhaps the exemplars. An iPhone learns the biometric of the phone's owner and stores this securely within the phone. Indeed, Apple (and Google) are quite clear that the fingerprint (template) doesn't (and never needs to) leave the device. When the owner wishes to unlock their device, or authorise a payment, their live fingerprint (template) is checked against the template stored on the device. If they match, the appropriate action takes place.

Iris and face biometrics are particularly sensitive to the effects of lighting on the capture of the biometric image that creates the templates and, in many circumstances, these requirements mean that iris and face biometric authentication is not encouraged. Aadhaar, for example, uses iris for uniqueness checking but (intends to use) one-to-one checking of fingerprints for authentication.

## 12.4 Where should privacy interventions be targeted?

Lawrence Lessig's work on the regulation of cyberspace[231] introduces the role that architecture and, particularly, code can play in regulating behaviour. He adds these to more traditionally understood regulatory mechanisms of laws, norms and markets. Code is particularly important, he suggests, because it provides anterior controls on action: the technology may prevent you from printing a protected document (or may prevent privacy risks by encrypting personal data). This can be done, even if the law says that, as the author of the work, you should be entitled to print the document or if a later court case determines that you were not entitled to do so.

Technological mechanisms can therefore provide some of the most restrictive and effective privacy protections for digital identity systems. Whether through the use of encryption of data, one-way generated sector specific identifiers, computationally enabled zero knowledge proofs, well designed technological systems ("privacy by design") can help minimise the privacy risks associated with identity systems.

Market issues that affect identity systems are described in more detail below.

The role of norms in Lessig's model is perhaps best understood in association with the role of levels of assurance in identity systems. Norms around relevant, effective and usable identity checking are likely to affect public trust in any resulting identity system. Any attempt to develop an emerging norm that requires extensive identity checking for even trivial identity checks is likely to be resisted, whereas norms of limited identity checks (or mechanisms (cash) to bypass some checks) may prove difficult to change.

Finally, laws such as data protection legislation or softer forms of regulation such as trust frameworks and contracts typically provide posterior controls. That is, whilst the law or trust framework may not permit certain privacy invasive activities (or require privacy impact assessments before systems are implemented), in practice they are most effective in punishing transgressions from these requirements, punishments that occur after the privacy damage has been done, as is the case with the South Korea identity breach.

### 12.4.1 What are the requirements around identity identifiers?

Most database systems, such as identity databases, use keys to uniquely distinguish one set of records from another. An identity system identifier can therefore be used as a database key to enable the system to look up the records associated with that identifier. Thus, a tax record database might use the identity identifier as a database key for the tax records of that individual. When an individual presents their identifier, this can allow easy lookup of the associated database record (e.g. tax details). In many countries, the same identifier, e.g. US social security number, is used for many different databases and purposes. In others, such as the Austrian system, different (although cryptographically linked) identifiers are used for different sectors making it more difficult to look up all the records associated with one person across these sectors.

There are two types of identifiers, the totally random number and the "smart" number. Smart numbers are used in a number of countries and might include details of the person's date of birth, gender and a code for the place of birth. Although there are some useful data entry checks that can be undertaken to ensure that a smart number is of the correct format, there are

[231] E.g. Lessig, L. 1999. Code and other laws of cyberspace, New York: Basic Books.

also clear privacy issues with smart numbers, particularly if a person's date of birth is used as part of the security checks for other services. Questions of gender (at birth) are increasingly complex and again, undue privacy concerns arise if the identity credential appears to indicate a different gender to that of the person presenting the credential[232]. Coding a place of birth adds complexity if the identity number is to be used for non–citizens as well (e.g. international students) as there may be limited codes for the rest of the world. Further problems arise if the number space is insufficiently large as rewriting all systems that use it can be costly[233].

## 12.5 Who will pay for the identity system?

If identity credentials are to become a key infrastructure for a society, then important questions of how they are to be paid for arise. There are different models of charging for infrastructure provision that can be drawn upon, for example charging on a per use basis or top-slicing the costs and providing a standard provision for all. Choosing the right payment model can be problematic whether the identity provider is a government agency (particularly if it is required to be revenue neutral) or a commercial body.

Enrolment is one activity where costs can be clearly understood but this is likely to be a one-time only activity and might be compressed into a relatively short time period. Even allowing for individuals coming of age and new residents, once the population has been issued with an identity credential the only money to be made will be at renewal time.

There are important questions about peak loading associated with the cost of issuing credentials. For example, the previous UK proposals for a national identity card explicitly decided to link identity card issuance with passport renewal, thus ensuring a relatively stable enrolment process and a smooth associated renewal process, as approximately 1 in 10 of all passports need to be renewed each year. India, in contrast, has focussed on enrolling the majority of the population in a very short period of time and hence will have to reduce enrolment capacity (whilst maintaining universality) once the population is enrolled.

Paying for use of the credential is a useful alternative but has a tendency to both encourage unnecessary formal verifications (perhaps beyond the levels of assurance required by the risk assessment) as the identity provider is being paid per verification and discouraged by the body that is covering the costs of these verifications. Moreover, if the relying party is paying a fee per verification, it is likely that this cost will be passed onto the customer. This model also provides little incentive for more privacy friendly zero knowledge transactions.

Other market related concerns can also result in systems that are more privacy invasive. For example, the commercial considerations associated with implementing and running identity systems may result in business models that seek to monetise the identity data, for example, by using it to support targeted advertising, whilst discouraging suitable identity checking that might exclude potentially valuable customers.

### 12.5.1 Why questions of liability must be addressed?

An identity system provides assurance that claims about identities or attributes can, to a specified level, be supported. This means that the consumer of the claimed identity (the relying

[232] Currah P and Mulqueen T (2011) Securitizing gender: Identity, biometrics and transgender bodies at the airport. Social research 78(2), 557-582.
[233] Eriksson O and Agerfalk P (2010) Rethinking the meaning of identifiers in information infrastructures. *Journal of the Association for Information Systems* 11(8), 433-454.

party) should act based on the supported claim. The question that arises, however, is what happens if, for whatever reason, there is a problem with the identity claim (and hence the assurance given for it). If the relying party cannot rely on the claims being made (and the context is one where there are significant penalties associated with inappropriate actions) then this will encourage the relying party from supplementing identity system with its own system and procedures. It is likely, however, that these supplementary systems will not be as carefully designed as the underlying identity system and privacy and security risks are therefore likely to be higher.

Relying parties should not therefore be held liable for actions based on properly authenticated identity claims. As noted above, this does not necessarily mean the creation of an audit trail of transactions. What then of the liability of the identity providers. Here the complexity of the liability model grows as benefits and risks are shared unequally. In extremis, the identity provider privatises the sum of the benefits (e.g. payments for authentications) but socialises the risks (e.g. complete failure of trust in the identity system as a whole).

### 12.5.2 Is there a role for compulsion?

For countries introducing new identity credentials, questions of consent and compulsion become particularly significant from a market and rights perspective. They may cause significant disruption to the roll out of system. In such cases it is frequently stated that the new identity system is voluntary, not compulsory and that individuals can always choose not to have an identity credential.

If the credential is not compulsory, however, there is a classic chicken and egg problem associated with generating a critical mass of issued credentials. If there is nowhere that an individual can use their new identity credential more easily than existing documentation (because organisations may not bother to redesign their systems and work processes until there is a significant proportion of the user base holding the new credential) then why should that person bother to enrol with the new system?

As the critical mass of credential holders develops, then effective compulsion can arise. This might simply be a consequence of staff becoming more familiar with dealing with the new identity credential and that using alternative documentation, whilst not being unsuitable, simply results in a more cumbersome process. This problem might be particularly acute when comparing the speed of electronic authentication with manual checking of an analogue identity credential. In other cases, systems might be used that require details of the (voluntary) identity credential to be entered even though there is no legal requirement to have one[234].

Making the identity credential compulsory requires popular and political support and thus requires a series of practical uses that the credential can be used for. Ideally, these are not artificially created just to “generate demand” (solutions looking for problems) and are based on a realistic risk assessment. However, evidence from Europe suggests that the various electronic identity cards are used infrequently because most people have infrequent access to public services and those that do have more frequent access rarely need to formally identify themselves each time.

[234] http://www.caravanmagazine.in/vantage/how-get-married-without-aadhaar-number

# APPENDIX A CASE STUDIES

## A.1 Austria

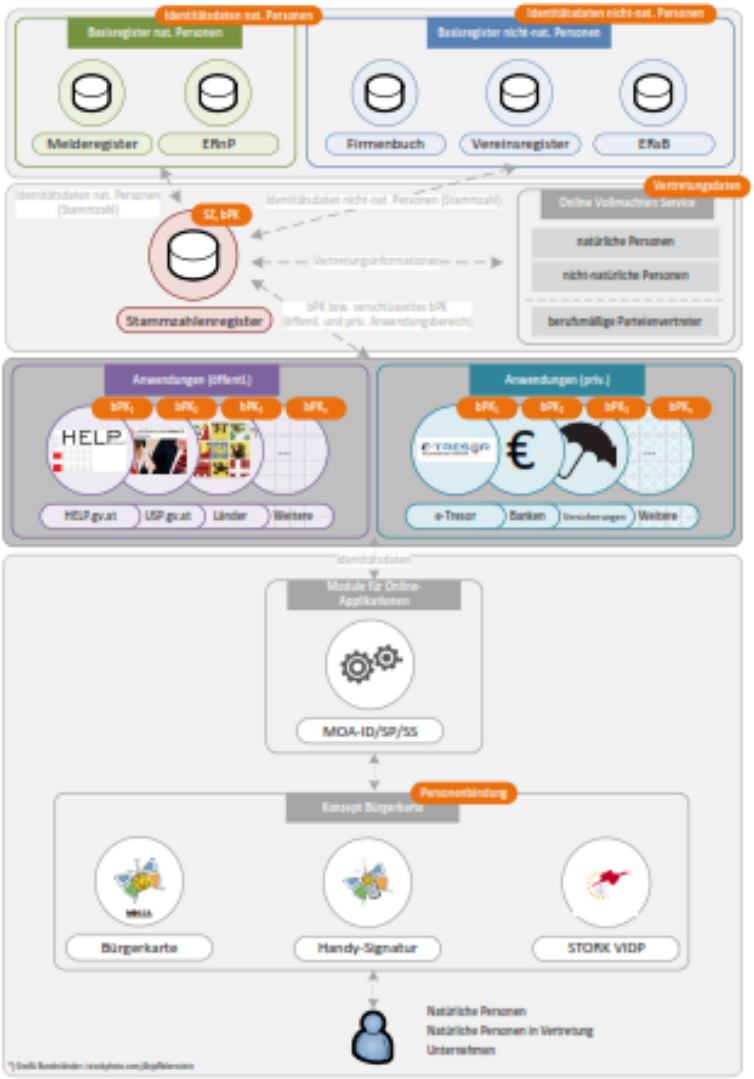
| | |
|---|---|
| **Description of Scheme** | High quality scheme in operation for over a decade. Privacy features, including separation of identities by sector, implemented through the issuance of sector-specific identities derived from a base CCR number. Available in both card and mobile formats. Uptake of card based ID effectively static, but mobile ID increasingly adopted by citizens. Received the United Nations Public Service Award 2014.[235] |
| **Links** | http://www.buergerkarte.at/ |
| **How identity is established** | Photo ID including:<br>• International passport<br>• Driving licence<br>• Austrian, German, Swiss or Liechtenstein ID card<br>• Professional affiliation cards (e.g. Chemist, Lawyer, Student, Military) |
| **Data collected, held, processed, and shared by the eID system** | • CCR number<br>• Name<br>• Date of birth |
| **How the eID integrates with other services** | The 12 services for citizens are:<br>• Income taxes: declaration, notification of assessment<br>• Job search services by labour offices<br>• Social security benefits<br>• Personal documents: passport and driver's licence<br>• Car registration (new, used, imported cars)<br>• Application for building permission<br>• Declaration to the police (e.g. in case of theft)<br>• Public libraries (availability of catalogues, search tools)<br>• Certificates (birth and marriage): request and delivery<br>• Enrolment in higher education/university<br>• Announcement of moving (change of address)<br>• Health related services (interactive advice on the availability of services in different hospitals; appointments for hospitals). |
| **Statistics on usage** | 2.5m cards, 420k mobile sigs, implemented on bank cards but discontinued due to minimal take-up, population 8.5m |
| **Positive press** | Very strong privacy by design from the outset and international co-operation. |

[235] https://joinup.ec.europa.eu/sites/default/files/egov_in_austria_-_january_2015_-_v_18_0_final.pdf

| | |
|---|---|
| Negative press | Some documented exploits and concerns that there are relatively few people with the expertise to be able to evaluate the system. Progressive approach to cryptography does not fall within the standard expectations of widely implemented systems. Also some general concerns regarding the security of card readers. |

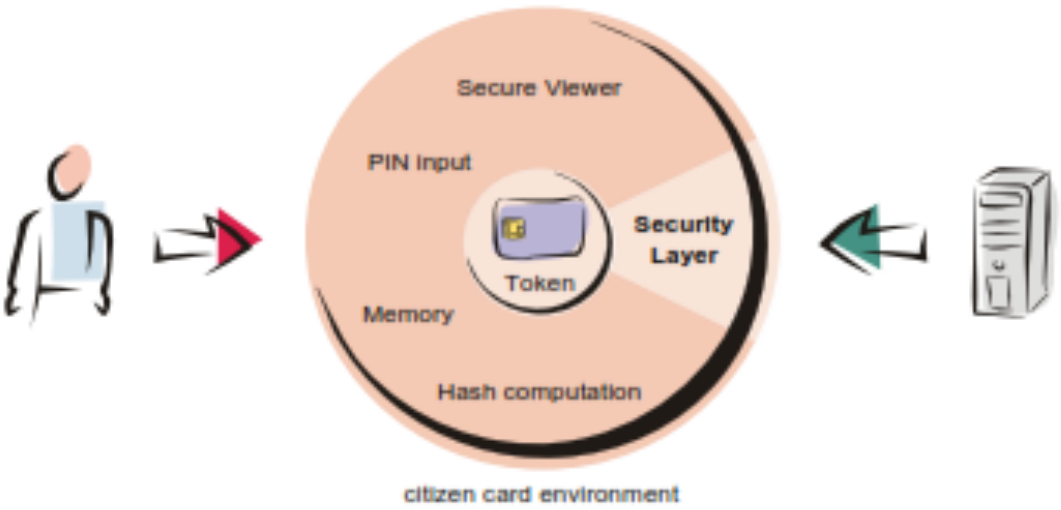
### A.1.1 Austria eID diagrams

**Figure 23 Big Picture eID**



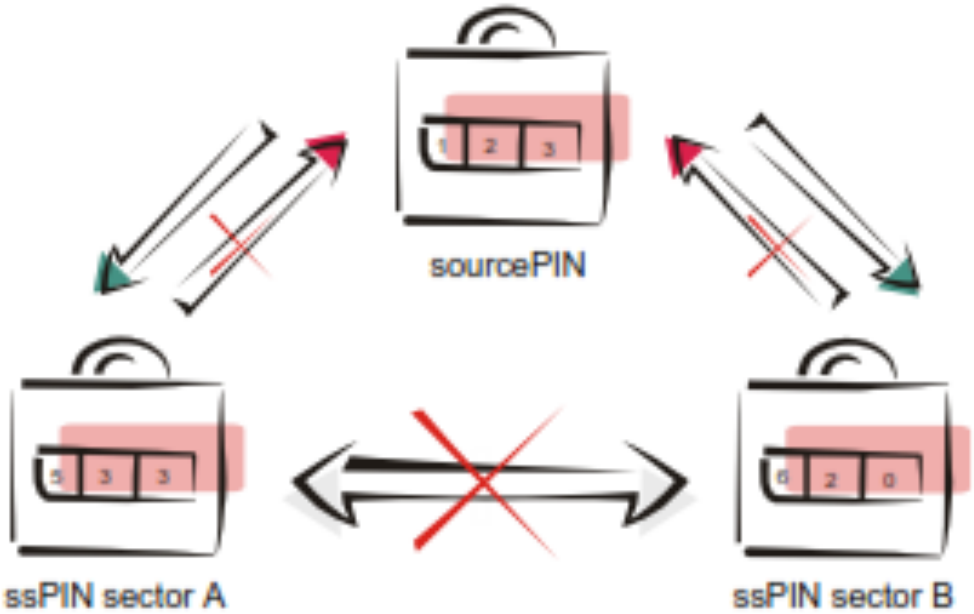

Source: Digital Austria, Federal Chancellery

Figure 24 Citizen card environment and token




Source: Digital Austria, Federal Chancellery

Figure 25 Identifiers derived for separate sectors




Source: Digital Austria, Federal Chancellery

[236]

[236] http://www.digitales.oesterreich.gv.at/DocView.axd?CobId=56936

## A.2 Canada

| | |
|---|---|
| **Description of Scheme** | Brokered CSP service allowing citizens to log onto digital government services using either:<br>• Banking credentials, which can be either online banking credentials or a contactless EMV payment transaction as a proxy identity performed against a contactless reader either embedded or attached to the citizen's PC.<br>• Government issued "GCKey" credential<br>The province of British Columbia has been issuing a public services based on the same EMV technology supported by the brokered CSP service. |
| **Links** | http://www.servicecanada.gc.ca/eng/online/mysca_credential_faq.shtml<br>http://securekeyconcierge.com/<br>http://www.cic.gc.ca/english/e-services/mycic.asp |
| **How identity is established** | The service provides authentication credential brokerage rather than digital identity per se.<br>Each service provider is required to establish the identity of the citizen accessing its service which can then be bound to the service provider-specific persistent identifier that is established. |
| **Data collected, held, processed, and shared by the eID system** | No explicit attribute data collected or shared by the service.<br>The authentication hub provides a separation of identifiers reducing the potential for linkability tracking. The identifier passed from the credential service provider to the hub is different from the identifier that is then presented to the service provider. Each service provider receives a different identifier for the same customer.<br>The hub in theory has the ability to track identifiers and usage but never gets to see attribute data, In the future if the hub is extended to support attribute exchange this may not remain true depending on the implementation. |
| **How the eID integrates with other services** | Service integrates with citizen-facing web services provided by government, using redirection within the web session, similar to federated identity, although the federated is moderated by the hub. |
| **Statistics on usage** | Over 6 million users (for GCKey[237] and SecureKey Concierge[238])<br>Over 120 services[239] |
| **Positive press** | http://www.huffingtonpost.ca/cleo-hamel/my-account-cra_b_4891673.html<br>http://www.reuters.com/article/securekey-technologies-idUSnBw065186a+100+BSW20140306 |
| **Negative press** | http://bccla.org/wp-content/uploads/2013/09/BC-Services-Card.pdf |

[237] http://www.newswire.ca/news-releases/gckey-service-exceeds-53-million-users---gckey-provides-individuals-and-businesses-with-electronic-credentials-for-securely-interacting-with-government-of-canada-online-services-516072801.html
[238] http://securekey.com/press-releases/securekey-concierge-service-surpasses-one-million-credential-milestone/
[239] http://securekeyconcierge.com/available-services/

## A.3 Chile

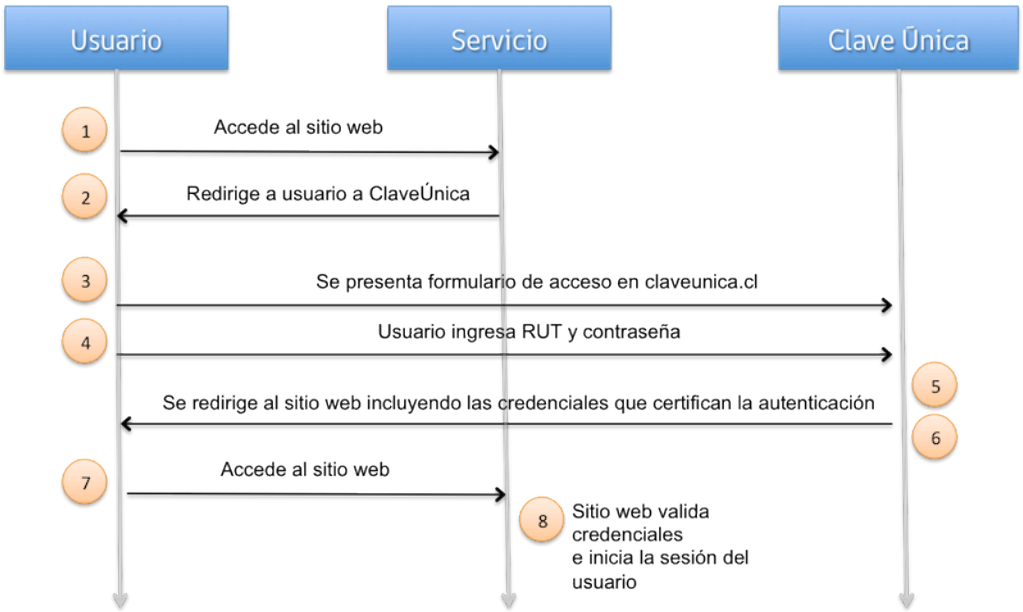
| | |
|---|---|
| **Description of Scheme** | Nascent scheme. Some teething troubles with deployment, but aiming to leverage OpenID Connect as part of the international iGov initiative. |
| **Links** | http://www.chileatiende.cl/ |
| **How identity is established** | Citizens<br>• RUN number<br>• Name, address and telephone number<br>Foreigners:<br>• Valid passport with residence visa stamped on it and certified by the registration of foreign international police.<br>• Photocopy of passport pages consisting in: identifying the holder, the validity of the passport and visa issued, the date of entry (if the visa consular), and foreign registration card PDI. |
| **Data collected, held, processed, and shared by the eID system** | • RUN number<br>• Digitised photograph, signature and fingerprint<br>• Name, address and telephone number |
| **How the eID integrates with other services** | Standard OpenID approach, via the claveunica.cl platform. Numerous relying parties in place. |
| **Statistics on usage** | At the same time they have been issuing OpenID 2 accounts for access to government and external services.<br>• 4M contactless Smart Cards<br>• 8M OpenID accounts (they started before the smart cards)<br>Of the 8M accounts, approximately 1 Million are active.<br>They currently have 30 Relying parties and more waiting to be connected.<br>There are about 10M of the older cards without biometrics that will be reissued over the next three years. |
| **Positive press** | Card used to deliver services for young people.[240]<br>Mentioned with France and Argentina in industry press as country developing services based on OpenID Connect as part of iGov initiative. |
| **Negative press** | Issues with ongoing strikes over several weeks in October and November by employees of the civil registry over pay[241]<br>Issues authenticating to access the social housing register.[242] |

[240] http://elurbanorural.cl/67-de-los-jovenes-indica-haber-hablado-por-internet-con-alguien-que-no-conoce-directamente/

[241] http://www.allchile.net/chileforum/viewtopic.php?f=11&t=13170&start=24

[242] http://www.biobiochile.cl/2015/12/10/falla-en-sistema-de-registro-de-hogares-mantiene-en-incertidumbre-a-postulantes-de-becas.shtml

### A.3.1 Chile eID diagrams




[243]

[243] http://instituciones.chilesinpapeleo.cl/page/view/claveunica

## A.4 Ecuador

| | |
|---|---|
| **Description of Scheme** | Early days of a scheme, which is not currently well formed. Technically ambitious, but structures unclear. |
| **Links** | http://sdw2015.com/speakers/i/593/#.Vm7Tt0qLTIV<br>http://www.registrocivil.gob.ec/ |
| **How identity is established** | Citizens<br>• Birth certificate<br>• Marital Status Certificate (single, married, divorced)<br>Foreigners:<br>• Original certificate of registration and registration of foreign immigrants, issued by the Directorate of Immigration, Ministry of Foreign Affairs and Human Mobility.<br>• Original passport.<br>• Original valid visa. |
| **Data collected, held, processed, and shared by the eID system** | Identification number, fingerprint code, names of the owner, the date and place of birth, nationality, gender, marital status, place and date of issue, expiration date, photo, holder's signature, signature of the competent authority, blood type and whether or not an organ donor.[244] |
| **How the eID integrates with other services** | Integration unclear. Implementation still subject to ongoing legislation.<br>Digital currency also currently being implemented.[245] |
| **Statistics on usage** | Too early to tell. |
| **Positive press** | A sophisticated eID program with multi-application cards deployed to all citizens, with NXP's SmartMX platform of dual interface chips and applications including eGovernment, banking and public transport. The country of 15 million is deploying a number of digital public services to citizens, companies and organizations. In addition to traditional ID vetting use cases, Ecuador's eID cards enable citizens to travel inside the Andean Community, perform electronic signature operations, and access social benefit and welfare services provided by the Ecuadorian government.[246] |
| **Negative press** | Some coverage following a press release by NXP in November 2014. Very little else since then.[247]<br>Digital environment politically turbulent in Ecuador.[248] |

[244] http://www.elmercurio.com.ec/507253-la-nueva-cedula-dni-se-emitiria-luego-de-2-anos/

[245] http://www.theguardian.com/world/2015/feb/26/ecuador-digital-currency-dollar-rafael-correa

[246] http://www.secureidnews.com/news-item/reid-national-eid-series-different-americas/

[247] http://www.electronicspecifier.com/communications/ecuador-to-utilise-versatile-secure-eid-card-technology

[248] https://freedomhouse.org/report/freedom-net/2014/ecuador

## A.5 Estonia

| | |
|---|---|
| **Description of Scheme** | Leading national eID originally launched in 2002. Includes:<br>• eID smart card that includes digital certificates (and associated keys) that can be used for document signing.<br>• Mobile identity which allows the SIM to be used to perform digital signatures. It also acts as an electronic use only token (i.e. not for face-to-face). A key benefit of this is fast issuance and replacement allowing citizens to still access digital services whilst waiting (weeks) for a re-issued eID card.<br>• X-ROAD integration services allowing government services to access the eID. |
| **Links** | http://eid.eesti.ee/index.php/EID_application_guide<br>http://www.id.ee/?lang=en&id= |
| **How identity is established** | Every citizen over 15 required to have eID. Managed through a central registry. |
| **Data collected, held, processed, and shared by the eID system** | The eID smart card contains the same information as that embossed onto the front of the card: full name, date and place of birth, personal identification number, citizenship, gender, type of residence permit) as well as card device information (e.g. serial number).[249]<br>The mobile eID Mobile ID is just certificate on SIM. It appears identity data not present but TSP (Trust Service Provider) looks up certificate (from CA) from which basic identity information (name, personal identification number) can be extracted.[250] |
| **How the eID integrates with other services** | X-ROAD is integration layer that allows service providers to connect to TSP etc.<br>Service providers maintain their own data (i.e. anything beyond the basic details) |
| **Statistics on usage** | Ubiquitous in Estonia, widely integrated. |
| **Positive press** | https://www.secureidentityalliance.org/index.php/resources/preview?path=14-06-02-SIA-Estonia%2BVisit%2BReport.pdf<br>http://www.economist.com/news/international/21605923-national-identity-scheme-goes-global-estonia-takes-plunge |
| **Negative press** | http://balticbusinessnews.com/article/2015/10/27/fault-in-id-cards-quickly-becoming-a-nightmare-for-estonian-government |

## A.6 Kenya

| | |
|---|---|
| **Description of Scheme** | National eID scheme about to launch |
| **Links** | https://www.ecitizen.go.ke/ |

[249] https://eid.eesti.ee/index.php/General_information_for_developers#Using_the_personal_data_file
[250] https://eid.eesti.ee/index.php/A_Short_Introduction_to_eID#Mobile_ID

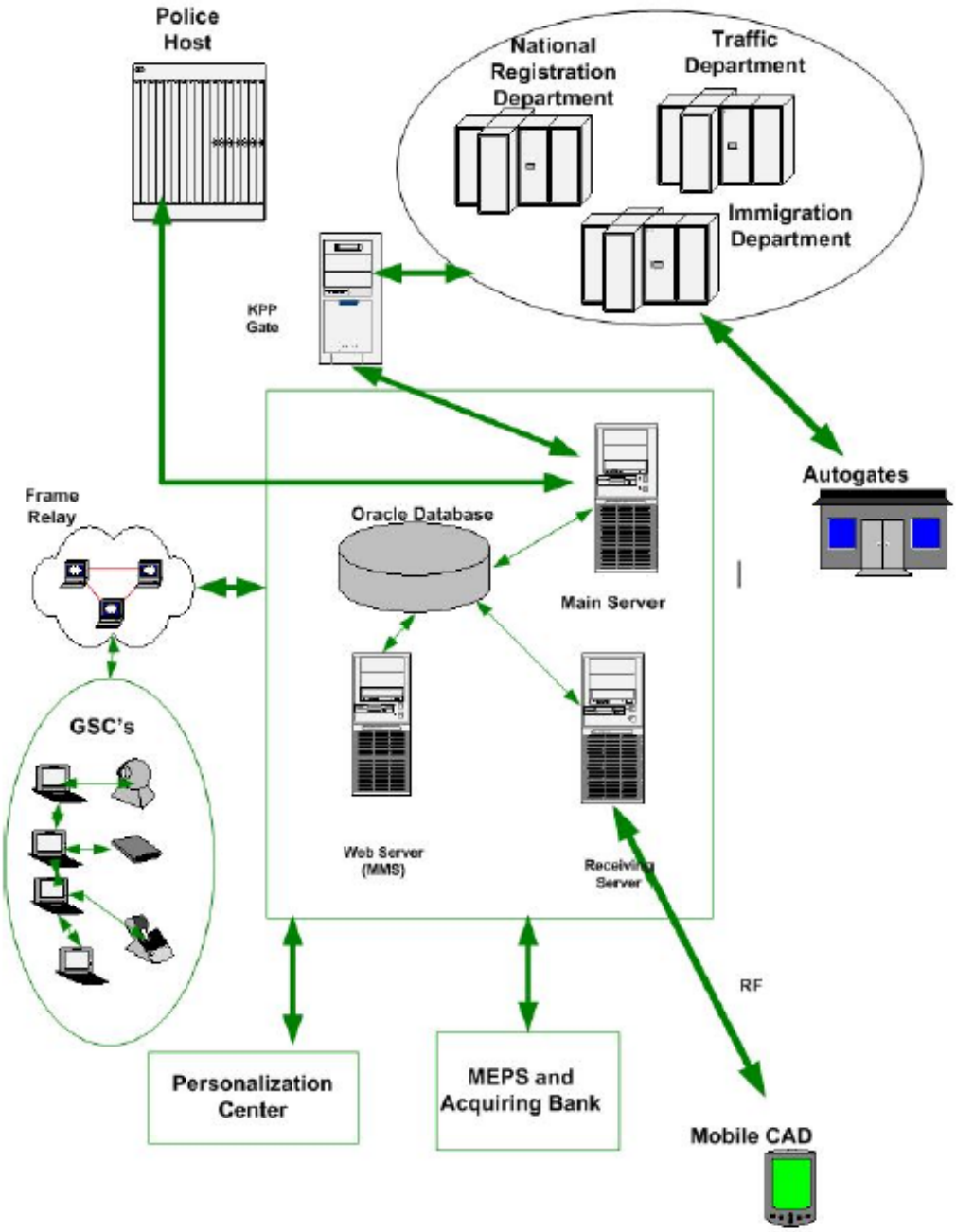
| | |
|---|---|
| **How identity is established** | • Copy of Birth Certificate of the applicant<br>• Copy of Valid Kenyan passport of the applicant<br>• Copy of Kenyan passport and/or Identity Card of either parent of the applicant<br>• The original 1st generation Identity card (Requirement for holders of 1st generation Identity card)<br>OR :<br>• Copy of Certificate of Registration<br>• Copy of Naturalization Certificate<br>• Copy of Valid Kenyan passport of the applicant<br>• The original 1st generation Identity card (Requirement for holders of 1st generation Identity card) |
| **Data collected, held, processed, and shared by the eID system** | • Name in full<br>• Gender<br>• Declared tribe or race<br>• Date of birth or apparent age, and place of birth<br>• Occupation, trade or employment<br>• Place of residence and postal address<br>• Finger and thumb impressions, and in case of missing fingers and thumbs, palm or toe and toe impressions<br>• Date of registration |
| **How the eID integrates with other services** | Information not yet available |
| **Statistics on usage** | Kenya was due to start issuing eID cards by October 2015, but as at December 2015, the process was yet to start. Public information was not readily available, the information provided below is based on the existing ID card. |
| **Positive press** | Too early for press reports |
| **Negative press** | Too early for press reports |

## A.7 Malaysia

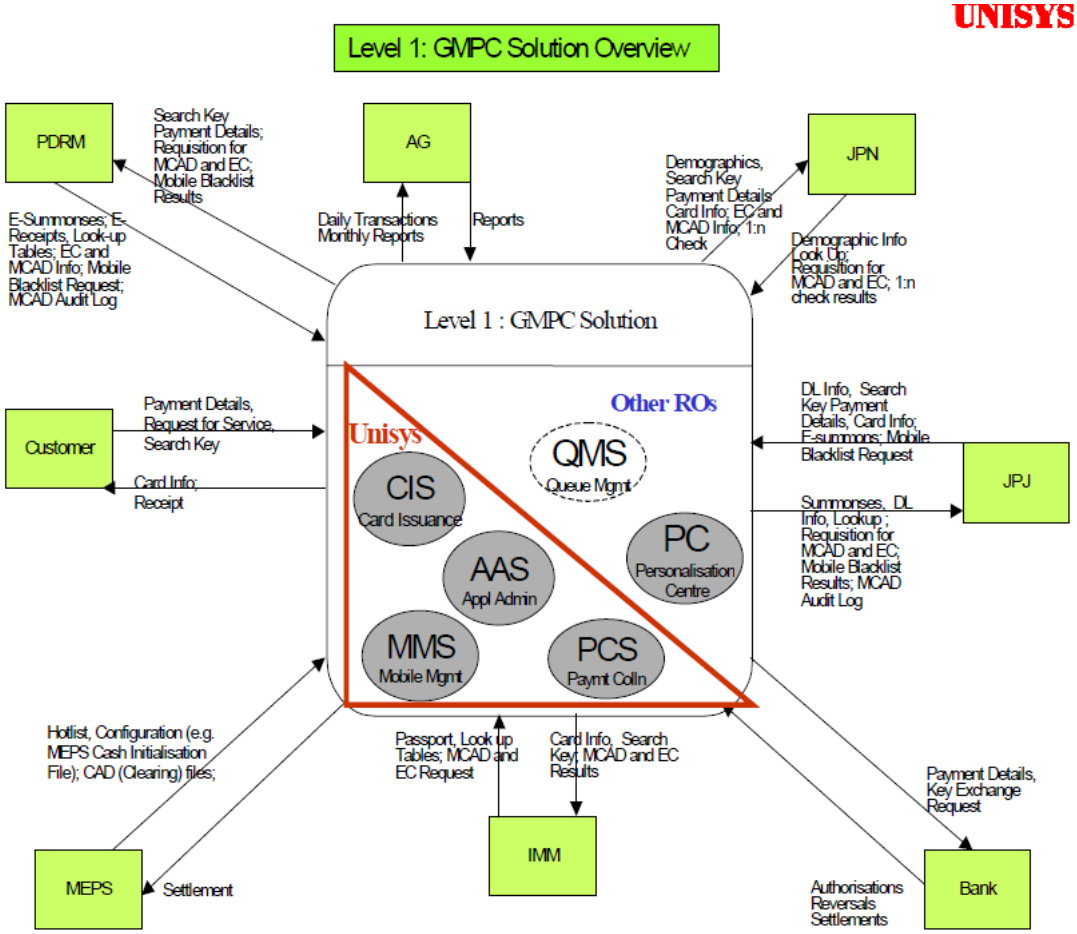
| | |
|---|---|
| **Description of Scheme** | Established eID scheme |
| **Links** | http://www.jpn.gov.my/en/informasi/tahap-keselamatan-mykad/# |
| **How identity is established** | Citizen<br>• Birth Certificate or Adoption Certificate or Citizenship Certificate or confirmation of citizenship status or Citizenship Form<br>• Identity card<br>• A guardian who is not a relative of the applicant must bring along the document confirming guardianship from the Department of Social Welfare<br>Foreigner<br>• Passport if the applicant is a non-resident<br>• Non-citizen sponsors must bring passports from their country of origin |
| **Data collected, held, processed, and shared by the eID system** | • Name<br>• Address<br>• Race<br>• Citizenship status<br>• Religion (for Muslims)<br>• Fingerprint minutiae |
| **How the eID integrates with other services** | Public domain system diagrams are provided below. From these it appears that the eID is integrated with numerous public services (e.g. transit, immigration, police) however it is not clear that there are clear boundaries between these services and their use of the eID to ensure good levels of privacy. |
| **Statistics on usage** | As at 2011, more than 96% of the population of 29.06 million[251] have MyKad[252]. 2015 population of Malaysia is 30 million[253]. |
| **Positive press** | Petrol subsidy:<br>• http://www.themalaysianinsider.com/malaysia/article/define-high-income-before-mykad-for-petrol-plan-urges-consumer-group-bernam<br>Multiple uses:<br>• http://www.funnymalaysia.net/10-benefits-of-having-a-mykad/ |
| **Negative press** | • PII can be easily deduced from the card[254]<br>• Uncontrolled access[255]<br>• PII stored on the card and available to anyone with the right card reader<br>• Fake MyKad in circulation[256],[257],[258] |

[251] http://www.tradingeconomics.com/malaysia/population
[252] http://www.ibimapublishing.com/journals/CIBIMA/2012/542549/542549.html
[253] http://www.livepopulation.com/country/malaysia.html
[254] https://en.wikipedia.org/wiki/Malaysian_identity_card
[255] http://www.malaysiakini.com/letters/25669
[256] http://www.coffeeticks.my/mykad-sold-for-between-rm8000-and-rm10000-each/
[257] http://www.themalaysianinsider.com/malaysia/article/revoke-all-identity-cards-issued-in-sabah-and-start-over-says-kitingan
[258] https://www.schneier.com/blog/archives/2011/09/complex_electro.html

### A.7.1 Malaysia eID diagrams



UNISYS



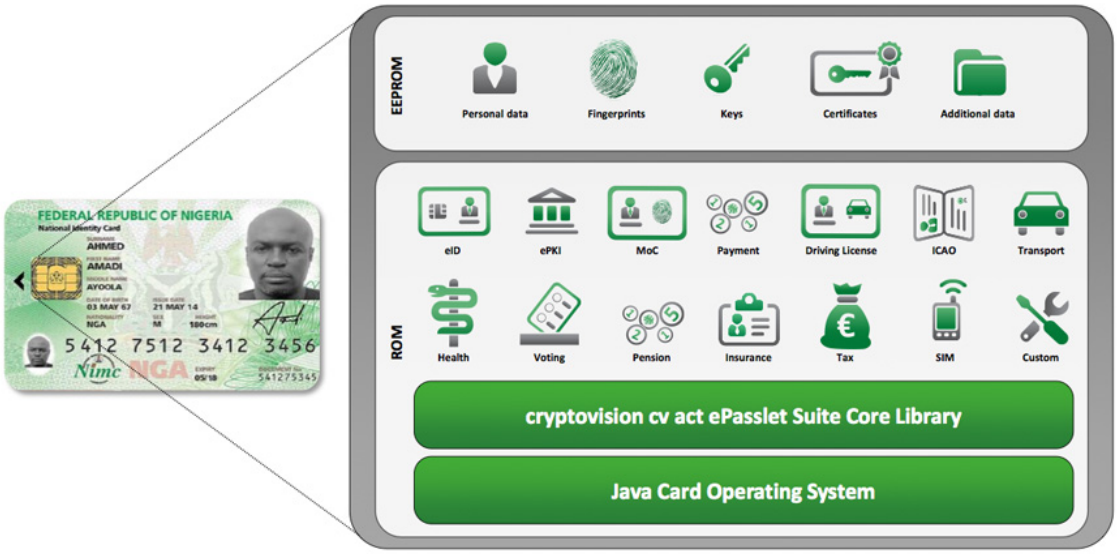

5 April, 2000 Unisys Confidential

### A.7.2 Nigeria

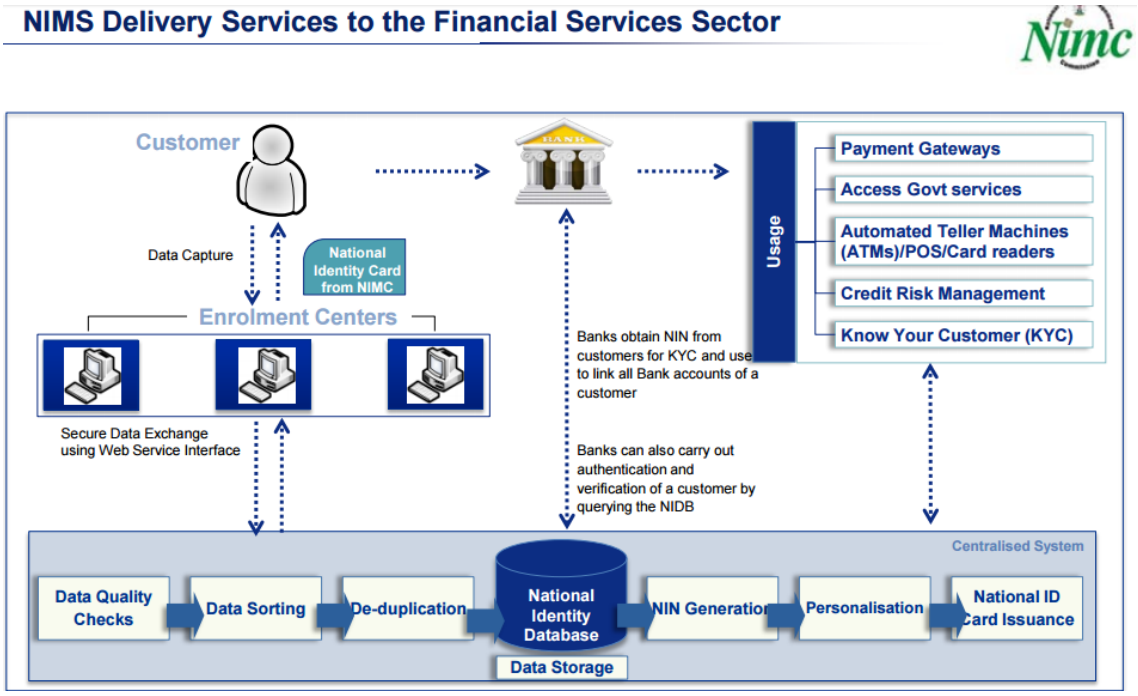
| | |
|---|---|
| **Description of Scheme** | National eID scheme based on Pakistan NADRA scheme |
| **Links** | http://www.nimc.gov.ng |
| **How identity is established** | • Old National ID card.<br>• Valid Driver's license.<br>• Valid International passport.<br>• Voter's ID card.<br>• Govt. staff ID card.<br>• State of origin certificate.<br>• Birth certificate/declaration of age |
| **Data collected, held, processed, and shared by the eID system** | • Personal Details<br>• Address Info<br>• Parent data<br>• Next of Kin Data<br>• Origin Data<br>• Identification documents Data<br>• Disability Data<br>• Signature<br>• Biometrics – facial and fingerprints |
| **How the eID integrates with other services** | Public domain system diagrams are provided below. From these it appears that the eID is integrated with numerous public and financial services (e.g. transit, health, pension, banking, etc) however it is not clear that there are clear boundaries between these services and their use of the eID to ensure good levels of privacy. |
| **Statistics on usage** | 15% (14,491,000)[259] out of a population of 185,921,240[260] |
| **Positive press** | None available |
| **Negative press** | • MasterCard's branding of the eID was controversial amid privacy concerns and the fear that foreign governments / organisations may have access to the data stored on the card[261] . NIMC refuted these concerns[262]<br>• Complaints about citizens not receiving their cards after enrolment and of service disruptions preventing enrolment both online and at centres[263].<br>• Citizens not collecting their cards[264]<br>• The Central Bank of Nigeria (CBN) was yet to direct banks to accept the card as a valid identification document for KYC, hence some banks were rejecting it for that purpose [265] |

[259] http://pubdocs.worldbank.org/pubdocs/publicdoc/2015/9/205641443451046211/ID4D-IntegrationAproachStudyComplete.pdf
[260] http://www.worldometers.info/world-population/nigeria-population/
[261] http://www.itwebafrica.com/ict-and-governance/265-nigeria/234218-privacy-international-critical-of-nigerian-eid-cards
[262] http://www.nimc.gov.ng/?q=facts-about-national-e-id-card
[263] http://www.nimc.gov.ng/?q=nimc-apologises-services-distruption-nin-enrolment-centres
[264] http://nannewsnigeria.com/nimc-says-66m-national-id-cards-ready-collection

## A.7.3 Nigeria eID diagrams




*Courtesy CryptoVision*[266]




Courtesy NIMC[267]

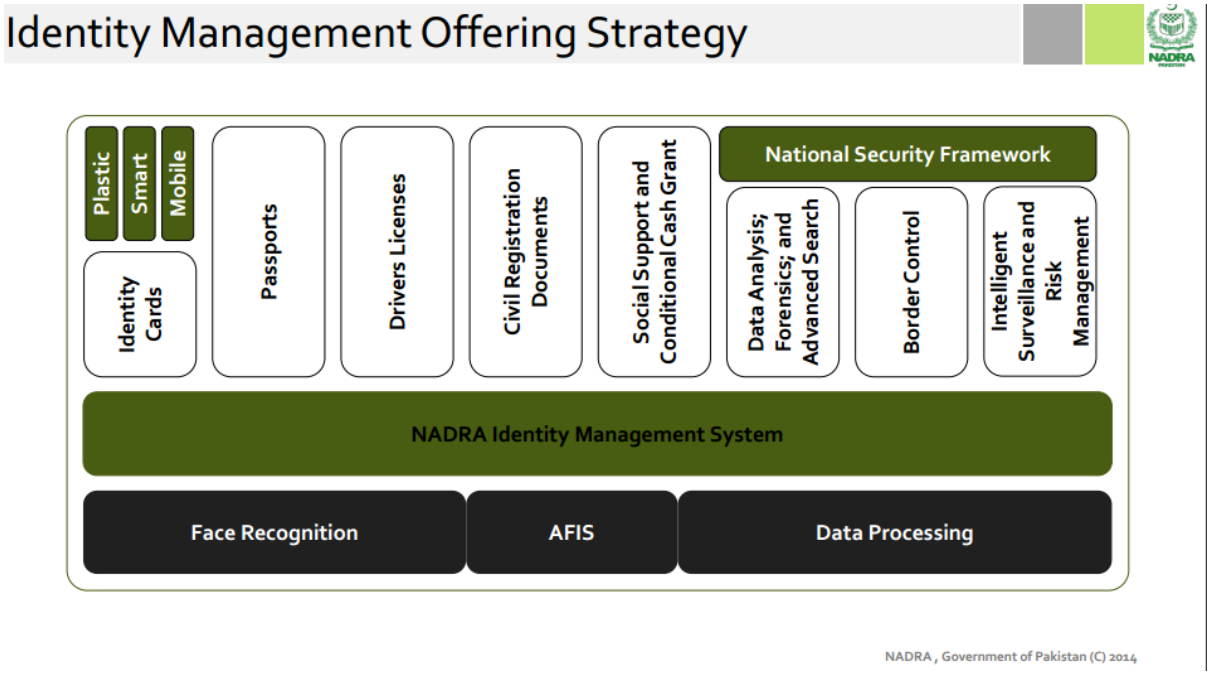
[265] http://nigerianreviews.com/banks-refuse-usage-of-new-national-identity-card/
[266] https://www.cryptovision.com/tag/nigerias-national-electronic-identity-card/
[267] http://www.nimc.gov.ng/sites/default/files/value_proposition_cashless_economy.pdf

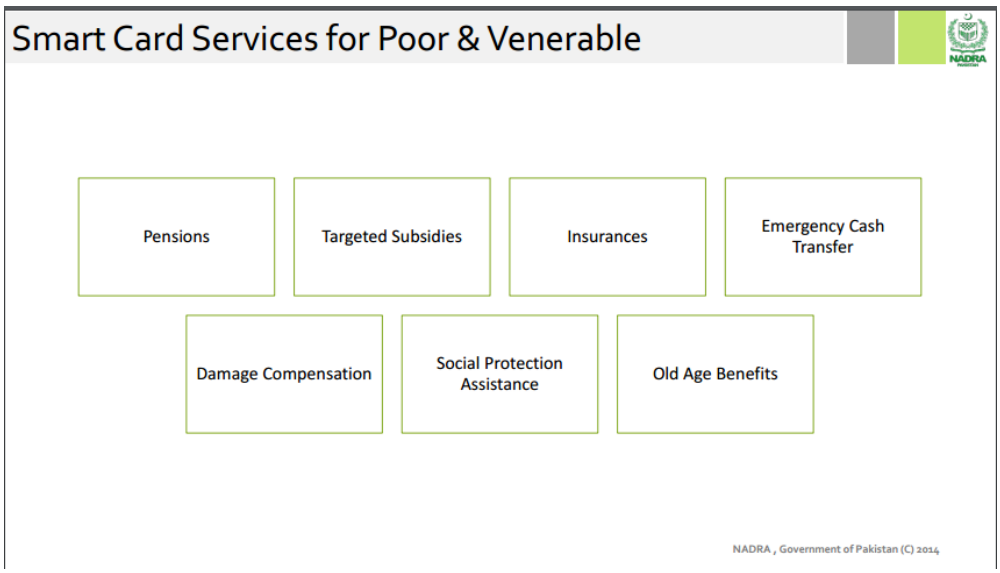
## A.8 Pakistan

| | |
|---|---|
| **Description of Scheme** | National eID scheme |
| **Links** | https://www.nadra.gov.pk/ |
| **How identity is established** | • Birth Certificate or<br>• Old NIC or<br>• Matriculation Certificate or<br>• CNICs of immediate/blood relatives<br>• Citizenship certificate issued by MOI |
| **Data collected, held, processed, and shared by the eID system** | • Legal Name<br>• Gender (male, female, or transgender)<br>• Father's name (Husband's name for married females)<br>• Identification Mark<br>• Date of Birth<br>• National Identity Card Number<br>• Family Tree ID Number<br>• Current Address<br>• Permanent Address<br>• Religion<br>• Signature<br>• Photo<br>• Fingerprint (Thumbprint) |
| **How the eID integrates with other services** | Public domain system diagrams are provided below. From these it appears that the eID is integrated with numerous public services (e.g. transit, immigration, police) however it is not clear that there are clear boundaries between these services and their use of the eID to ensure good levels of privacy. |
| **Statistics on usage** | |
| **Positive press** | • Cash Benefit to Internally Displaced Persons (IDP)s of Swat & Malakand<br>• Cash Benefit to 2010 Flash Flood Victims<br>• Cash Benefit to Ultra Poor & Vulnerable Population - The Benazir Income Support Program (BISP)<br>• World Bank report[268] |
| **Negative press** | • Identity fraud[269],[270]<br>• Issuing cards to non-citizens[271]<br>• Granting PII data access to companies instead of authenticating citizens[272] |

[268] http://www.worldbank.org/content/dam/Worldbank/Event/social-protection/Building_Robust_Identification_Systems_Session_Packet.pdf
[269] https://www.privacyinternational.org/node/334
[270] http://tribune.com.pk/story/941493/pakistan-probes-100-id-cards-for-militants-scam/
[271] http://www.thenews.com.pk/print/56205-nadra-and-fraud
[272] http://www.dawn.com/news/870657/unauthorised-sims-breach-of-security

### A.8.1 Pakistan eID diagrams




Courtesy World Bank[273]

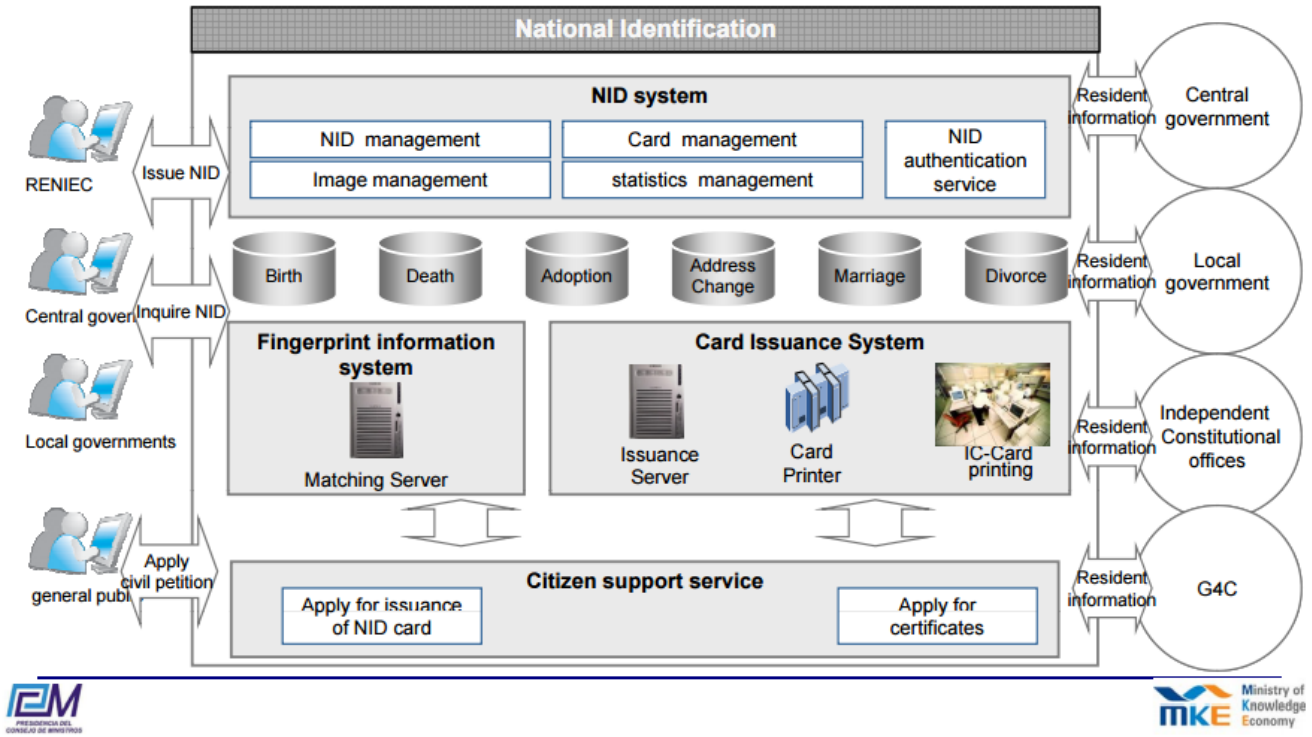
[273] http://www.worldbank.org/content/dam/Worldbank/Event/social-protection/Building_Robust_Identification_Systems_Session_Packet.pdf

## A.9 Peru

| | |
|---|---|
| **Description of Scheme** | National eID scheme |
| **Links** | http://www.reniec.gob.pe/portal/homeDepartamento.htm |
| **How identity is established** | • Original birth certificate |
| **Data collected, held, processed, and shared by the eID system** | • Given names and surnames,<br>• Unique identification number,<br>• Date of birth<br>• Marital status<br>• Photograph<br>• Fingerprint<br>• Voting number |
| **How the eID integrates with other services** | |
| **Statistics on usage** | In 2016, the current DNI expires, so thereafter, the delivery will be more intensive. In that sense, it is expected that by 2021 the coverage reaches 79% of the population. |
| **Positive press** | Praised in Latin American ID circles[274]<br>Currently small scale but will need to ramp up significantly as existing DNI cards expire in 2016.<br>'Reniec Facial Mobile' is a govt. android app that allows the user to make changes to housing records and also organ donor status on passport, matching the photograph taken on the mobile with one held in the RENIEC database. |
| **Negative press** | Some complaints regarding the price, as this could exclude poorer citizens<br>Also concerns regarding insufficient registration of births (missing birth certificates). |

[274] http://www.andina.com.pe/Ingles/Inicio.aspx/movies/noticia-peru%E2%80%99s-electronic-id-card-recognized-as-best-in-latin-america-562683.aspx

## A.9.1 Peru eID diagrams



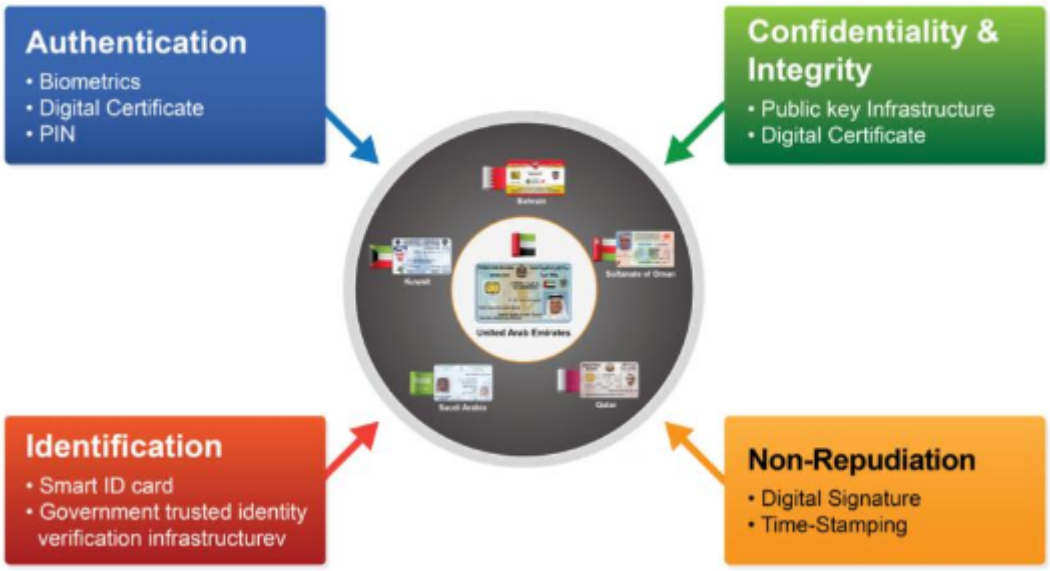

Courtesy[275]

[275] http://www.ongei.gob.pe/pdf/egovernment.pdf

## A.10 Saudi Arabia

| | |
|---|---|
| **Description of Scheme** | Establish national eID |
| **Links** | http://www.moi.gov.sa |
| **How identity is established** | • National ID or for citizens<br>• Residence permit for foreigners |
| **Data collected, held, processed, and shared by the eID system** | • National ID or<br>• Residence permit<br>• E-mail address<br>• Mobile number |
| **How the eID integrates with other services** | Fairly standard PKI with smart cards, as implemented across six GCC countries.[276] |
| **Statistics on usage** | |
| **Positive press** | Significant up-take both online and in person, with new registration centres being opened specifically for women. |
| **Negative press** | Very strong focus on security and management of residents.<br>All women (Saudi or ex-pat) required to have fingerprint ID by 2019. Original aim was to ensure a check all non-Saudi women as they entered and left the country. |

### A.10.1 Saudi eID diagrams




[276] https://www.paci.gov.kw/pdf/PACI_CP_V1.1.pdf

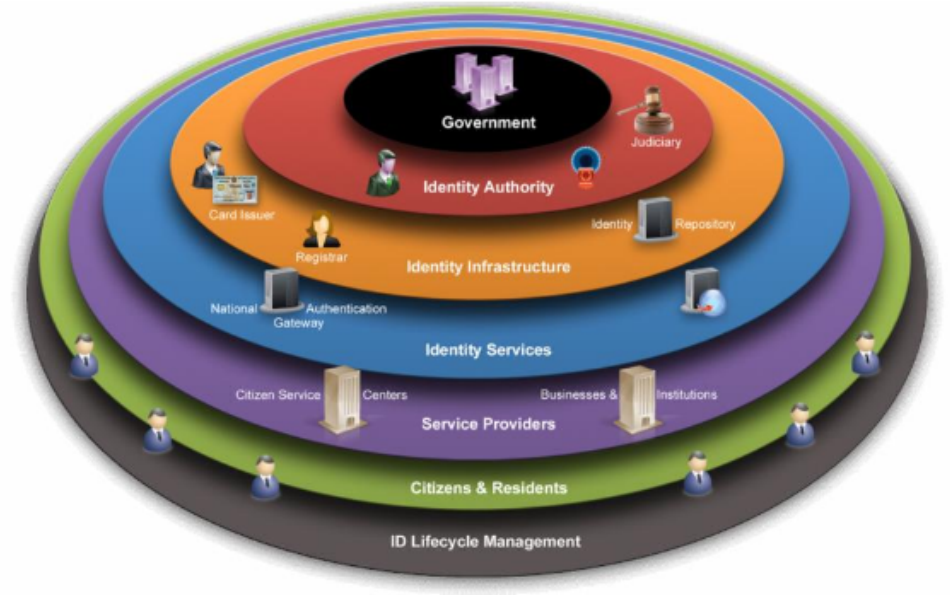



Figure 5. National Digital Identity—the GCC Context

277

## A.11 UK

| | |
|---|---|
| **Description of Scheme** | Brokered IDP scheme involving:<br>• Government built hub<br>• A range of accredited and to be accredited private sector identity providers.<br>When accessing a government site citizens are given a choice of identity provider, who, on first use, take the citizen through an identification process and on subsequent usage authenticate already identified citizen. A core set of attribute data (name, address, date of birth, gender) is then shared with the service provider.<br>The objective of the scheme is the create a market place for digital identity such that government can share the costs of identity with the private sector. |
| **Links** | https://gds.blog.gov.uk/category/id-assurance/ |
| **How identity is established** | Identity providers establish the identity of citizens following government specified standards and guidelines using a combination of “citizen”, “money” and “life” sources. This will typically include passport or driving license documents, performing an identity check with a credit reference agency, seeking other corroborating evidence such a utility bills and checking databases of known compromised identities. |

277 http://www.ijpis.net/ojs/index.php/IJPIS/article/view/113/pdf

| | |
|---|---|
| **Data collected, held, processed, and shared by the eID system** | For the government digital identity services IDPs store the core data attributes established in the identification process and transactional history associated with the citizen's use of the IDPs service. The IDP has no visibility of the service provider being accessed by the citizen.<br>All transactions flow through the hub and the core data attributes are carried in every transaction. The hub is stateless and therefore if operated correctly should not store or aggregate any data. |
| **How the eID integrates with other services** | Service integrates with citizen-facing web services provided by government, using redirection within the web session, similar to federated identity, although the federation is moderated by the hub. |
| **Statistics on usage** | Currently still in public beta testing. The tax self assessment deadline (January 2016) is expected to lift the number of users from a few hundred thousand to over 2 million,<br>https://identityassurance.blog.gov.uk/2015/05/14/the-next-6-months-services-that-plan-to-start-using-gov-uk-verify-2/ |
| **Positive press** | http://www.computerweekly.com/news/4500250478/GDS-targets-three-million-new-users-of-Govuk-Verify-service-in-next-12-months |
| **Negative press** | http://www.computing.co.uk/ctg/news/2414194/govuk-verify-identity-management-system-riddled-with-severe-privacy-and-security-problems-warn-ucl-academics<br>http://www.computerworlduk.com/data/govuk-verify-explained-3626323/ |

# APPENDIX B CROSS REFERENCE

Below, we cross reference the contents of this report with the questions posed in the original RFP.

## B.1 Privacy

- What is the impact of digital identity systems on individual privacy rights?
  - How might digital identity systems protect, infringe upon, or reshape individual privacy rights?
  - What are the market forces shaping privacy? Are there other relevant forces to consider (e.g. voter behaviour, incumbent interests, etc.)
  - What is the likely impact of various leading private sector identities (e.g., Facebook, MasterCard smart cards, mobile phones, etc.) on privacy
  - Does the impact on privacy differ for single centralized digital identity systems versus multiple fragmented systems? Does it differ for foundational (e.g. general IDs like Aadhaar) versus functional systems (e.g. electoral IDs)

As a result of technological advancements, modern understandings of individual privacy rights focus principally on the protection of personal information, and the ability of an individual to control who holds what data about them (section 2.2). At the same time, we have seen the expansion of the definition of personal information, as the volume of data digitally created grows exponentially each year (section 3). In addition to constitutional and international human rights law, a body of law has emerged to regulate the privacy of personal information, generally known as “data protection law”. The European Union’s own approach to data protection law has been incredibly influential in shaping the evolution of data protection law in countries outside of Europe (section 3.3). In addition, in recent years, and in particular since the Snowden revelations, greater customer demand has emerged for “privacy-positive” products, and a number of major technology companies have responded by rolling out products which minimise the impact on individual privacy (section 2.3). At the same time, there is a growing trend towards the mandatory use of identity, particularly online, which not threatens to undermine privacy and impede anonymity (section 3.4.3), but also may result in exclusion of those who don’t have access to identity.

The centralised generation, storage, and retention of personal information, and its transmission, creates numerous and often serious risks to individual privacy rights. For this reason, digital identity systems, depending on their design and deployment, will inevitably create some concerns from the perspective of privacy, which concerns can be mitigated through regulatory, technical and corporate factors (section 5).

It is possible to identify seven architectural models of digital identity technologies currently available or being debated: monolithic internet providers, such as Facebook Connect; Federated internet identity providers, such as Open ID Connect; State-issued eID cards; Brokered Identity Providers, such as the UK’s Verify programme; Brokered Credential Service Providers, such as the Canadian system; Personal Identity Providers, such as MyDex; and No Identity Provider-based systems, such as Bitcoin (section 6).

The extent to which a particular digital identity system impacts positively or negatively on privacy will depend not only on the architectural model of the system, but will reflect the particular context the system is designed and deployed in. To this end, three sets of factors are relevant to a system's impact on privacy: the regulatory context, the technological aspects, and the commercial viability of the system (section 5).

With respect to the regulatory context, two clear indicators exist as to whether a system can be conductive to protecting, rather than endangering, privacy: the existence of a comprehensive data protection law, and the existence of a specialised independent data protection authority with adequate human and financial resources (section 5.1.2).

The choice of technologies employed in the delivery of digital identity systems will also be critical to assessing the privacy impacts of a particular system (section 5.1.3). While some technologies are inherently privacy-friendly, others will have positive or negative impacts on privacy depending on whether they facilitate or possess a number of characteristics, including unlinkability, secure communication and minimal data disclosure, and minimal data stored on an eID (section 5.1.3).

It is finally important to consider the commercial model that underpins a digital identity system, since it is the viability of this model that will be key to its long term success if even a single element of the delivery of the scheme is reliant upon a private sector, profit-driven entity. It is important to question whether it is in the commercial interest of a participating private sector entity to compromise on individuals' privacy, and whether the model presents a fair deal for the citizen (section 5.1.4).

- Where digital identity systems might threaten individual privacy rights, what is the best way to mitigate these risks?
  - What are best in class national regulatory or legal frameworks (if any) that balance digital inclusion and privacy? What trade-offs (if any) are involved in using one framework vs. another?
  - Beyond the EU frameworks, what legal privacy approaches have emerging market governments developed that can be instructive examples (and why)?
  - What, if any, institutional frameworks, groups, or institutions can assist?
  - What, if any, are the relevant design and technical architecture principles?
  - Is it possible or desirable to create a lasting privacy framework given the rapid pace of change in the field?
  - What citizen recourse and transparency features can help mitigate the risks?
  - To what extent does a requirement that use of eIDs is mandatory affect individual privacy considerations?
- How can digital identity systems maintain privacy protections while fostering a rich ecosystem of digital identity applications?
  - Digital identity systems serve as platforms for uses ranging from mobile payments to health service provision to social benefit targeting. How can digital identity systems balance privacy safeguards and support for diverse and innovative uses? Compare different approaches to privacy and highlight their respective benefits and risks.

It is generally agreed that the European Union's Data Protection Directive (soon to be superseded by the General Data Protection Regulation) is the most rigorous regulatory regime pertaining to the protection of personal information in the world. Many other non-European States have emulated its approach. The only real competing regulatory approach to data protection is the United States' self-regulatory and sectoral regulation approach to privacy. While reasonable minds can differ on this point, it is posited here that the US will ultimately bring its regulatory regime into line with the European approach (section 3.3). Emerging economies have developed or deployed a different legal approach to the regulation of data protection.

In addition to the general public law framework applicable in a particular country, each identity system will have a legal basis and be governed by contractual agreements between the parties and any standards adopted by the parties (usually specified in the contract) (section 4.7).

It is possible to distill certain legal and technical principles which should guide any identity system in order to ensure that it is privacy-friendly (section 5). Such principles have been well-developed in Europe (the OECD Principles, Council of Europe's Convention 108, and the EU Data Protection Directive), the US (the Fair Information Practice Principles), and in the context of identity (the 7 laws of identity). We believe there is therefore general support and recognition that the relevant principles to privacy in the context of digital identity can be encapsulated in our Digital Identity Privacy Principles (DIPPs), as follows:

1. Fair and lawful processing – Identity schemes must be regulated by strong legal frameworks that require data to be processed fairly and lawfully
2. Adequacy and quality – Data is to be adequate, relevant and non-excessive in relation to the purpose for which it was collected, accurate and up to date
3. Explicit and legitimate purposes – Data is only to be collected and processed for specified, explicit and legitimate purposes
4. Minimal disclosure for a constrained use – Data is only to be used for ways which are compatible with specified purposes and disclosed to parties only to the extent it is strictly necessary
5. Openness and transparency – Individuals should be given the greatest amount of information possible about how their data is processed, used, stored, disclosed, retained and deleted
6. Individual ownership and control of data – individuals should have the ability to access information about what data is held about them, who has access to that data and on what conditions it is being processed. Individuals should be able to correct and transfer their data where applicable.
7. Accountability and auditing – there should be regular independent auditing of identity providers, and individuals should have avenues of redress if their data is misused or incorrectly disclosed.
8. Consent – identity schemes should always obtain the consent of individuals to use their data for the purposes specified.
9. Data minimization and avoidance of honeypots – Identity providers and others involved in the identity process should request and store the minimal amount of data necessary to perform their functions, minimising the creation of data honeypots at all times.
10. Sensitive data – there should be specific protections for sensitive data (that relating to political, religious, ethnic, or religious identity or health data).

11. Avoid exclusion – identity providers have an obligation to ensure that individuals' data is not used in a way that excludes them from access to services and opportunities.
12. Restrictions on transfer and disclosure of data – data should not be transferred to parties or locations outside of the data controller's direct control.
13. Pluralism and interoperability of systems and technologies – identity providers should choose systems and technologies that are useable, transferable and interoperable, and incorporate the individual.
14. Minimise the human element – The more a system is reliant on human intervention, the greater the potential for negligence and abuse.
15. Robustness of technology – The system incorporates strong cryptographic software and hardware.
16. Levels of assurance – The system requires and provides the highest levels of assurance of identity.

## B.2 Technology

- What are the most relevant technologies and technology architectures in the digital identity space, especially as they relate to privacy concerns?
  - What are the strengths and limitations of different technologies (e.g. card-based versus purely digital systems)?
  - Which technologies are most appropriate for which applications (e.g. financial services versus healthcare)?
  - What are the critical considerations for governments selecting digital identity technology (e.g. sophistication, downtime, batch vs. real time processing)?
  - How do different technologies support or undermine privacy and security?
  - Which architectures are most conducive to both effective operations and to protecting consumer privacy?

As noted above, the appropriateness of any one digital identity system and its impact on privacy will depend on a combination of regulatory, technical and commercial factors. Speaking generally, very few technologies employed in the delivery of digital identity today are *inherently* insecure, except if they are outdated or superseded technologies. Some technologies are less robust than others, however; we rank the strength of individual technologies in this report (section 5.1.3) and well as considering the overall risks to end-to-end systems (section 7). Financial services, especially the payments industry, has often been responsible for developing identity related technologies (especially through their use of smart cards and risk management engines). Whether these technologies are equally applicable in other applications or sectors will be a matter for the privacy impact assessment. One note to bear in mind is that banks often take a risk based approach to payments, tolerating certain levels of financial loss. In other sectors (e.g. healthcare) placing a financial value on a loss or other damages may be more difficult. A breach of privacy resulting in psychological trauma for example could be very difficult to quantify.

The scope of any digital identity service will determine the threats – who may attempt to undermine the service. The technology employed and how it is used will determine the

vulnerabilities – what weaknesses in the system may give the threat agent the opportunity to succeed. Together the threats and vulnerabilities produce risks (section 7).

As for any other digital service, the threats to a digital identity scheme fall into one of three broad classes:

- Threats to the confidentiality of the scheme. These are direct threats to the privacy of the enrolled participants, and to the service itself;
- Threats to the integrity of the scheme – that is, that the data held by the scheme is altered by an unauthorised third party;
- Threats to the availability of the scheme. These are threats to the service itself, either disrupting enrolment or the subsequent provision of identity-related services, such as identity assurance.

These threats might further arise through a number of avenues, such as:

- Malevolent threats (entities operating illegally to disrupt a digital identity service, to undermine its integrity, or to undermine the privacy of individuals).
- Legitimate businesses that potentially compromise the privacy of personal information through their legal business practices.
- Passive threats to the service, which arise through either incompetence or negligence on the part of those responsible for its provision.

The range of threats a digital identity service might be exposed to is summarised in the report (section 7.1), and rated as high, medium or low; these threats range from identity theft to mass surveillance, poor operational security to denial of service attacks. We also list types of vulnerability that may exist in a digital identity system, associating them with the DIPPs established above. For each type of vulnerability, we provide examples of different severities of vulnerability (high, medium and low severity) (section 7.2).

The privacy risks (threats + vulnerabilities) posed by digital identity schemes can be mitigated through practical and technical measures. Whether there is sufficient motivation to undertake such measures will depend on a combination of political, social and commercial pressures. In addition, it is important to recall that the parties that are required to mitigate risk may not be the parties with the interest in doing so. The internet giants again provide a clear example of this. Their primary interest is in maintaining advertising revenues, which come from advertisers not consumers. Furthermore, complex innovative systems can develop in unpredictable ways with unintended consequences. This is especially true where there is pressure to roll out new features and service at a rapid pace, meaning that there is little time for a measured consideration of potential privacy.

The key areas of impact and risk can be described as follows (section 7):

| | |
|---|---|
| Regulatory | • Inadequate data protection law<br>• Weak institutional mechanisms |

| | |
|---|---|
| Technical | • Identifiers that are linkable<br>• Unnecessary collection of data<br>• Unintended disclosure of data |
| Commercial | • Commercial models, especially those based on deriving additional value from user data, in conflict with privacy. |

In the report, we describe specific mitigation strategies for each digital identity model (section 8.4).

- What are the technological opportunities and limitations for establishing digital identities as a public good?
  - What are the technological issues around interoperability? Where can technologies be integrated versus where do they need to be implemented anew?
  - What are the technological issues around inclusion/exclusion?
  - What organizing forces in the technology space might establish these as public goods (e.g. civil society coalitions)? What organizing forces might erect opposition (e.g. the smartcard industry)

A number of other contextual, practical and financial issues will impact upon the establishment of digital identity systems, and their acceptance as a public good. These include (Section 11):

- Communications issues such as the coverage of data communications services, the need for online access to digital identity databases, which can be somewhat mitigated by offline capture of registration application details, and the need for online authentication; (section 11.1.1);
- The fact that that a significant proportion of emerging markets – though certainly not all – do not have data centres that are suitable for use in support of digital identity services, or have data centres without sufficient physical security; (section 11.1.2);
- The lack of a commercial case for a digital identity scheme (section 11.2;
- Fear of legal liability or uncertainty around liability considerations (section 11.3);
- The large scale of some digital identity schemes in emerging markets (section 11.4);
- Exclusion due to political capture and financial exclusion (section 11.5);
- Interoperability when digital identity schemes rely on a foundational identity scheme (section 11.6); and
- Funding, not least due to the scale of the initial registration task and the on-going operational costs (section 11.7), and related concerns about appropriateness (section 11.8).

We note that there is often a trade off between privacy and inclusion (section 7.4). Sometimes this is a deliberate choice on the part of the identity provider. Other times it is an implicit

consequence of the business model or architecture adopted by the identity provider (or scheme to which they belong). At one extreme the highly centralised monolithic Internet identity providers have built inclusive services that depend on free, simple and ubiquitous access to citizens, which has come at the cost of the privacy of the personal information of citizens. At the other extreme, highly decentralised and consumer-centric services high levels of privacy. These however suffer from issues related to inclusion, including the greater burden on the consumer to manage their own personal data and digital identity credentials.

Between these two extremes are models that offer some level of privacy and as well as some level of inclusion. The foundational national level schemes (including national eID schemes as well as the brokered digital systems) described in this document sit in this middle ground.

## B.3 Security

- What kind of protections can existing security systems provide (e.g. fraud-prevention algorithms, etc.)?
  - What (if any) are the best practices for digital identity security?
  - What (if any) are the points of debate?
- What are the primary threats?
  - How easy are existing systems to hack?
  - What are the potential consequences of hacking?
  - In the case of identity theft, what are the best practices for grievance redress
- What insights do existing use cases offer?
  - How do different actors—ranging from Google to India's UIDAI—think about security?
  - What lessons do past security breaches offer?
- What are the relevant design principles for safeguarding security in digital identity systems (e.g. federated versus non-federated architecture, data-sharing protocols, notifying consumers of use, and etc.)

This report treats security threats and vulnerabilities as inseparable from technical threats and vulnerabilities (sections 7 and 8). Adopting a mitigation strategy for threats and vulnerabilities will depend on the status of the digital identity system in question (section 8.2). If a digital identity system is to be developed from scratch that may provide the opportunity to design in privacy at the beginning. However, there are still many factors that will influence the overall approach taken to digital identity and as a result place constraints on the mitigation strategies that can be employed. Where an identity system is part of a regional system, enabling interoperability between digital identity systems can introduce new risks. Where digital identity systems already exist and are integrated into services addressing privacy issues may be difficult, due to the cost of making changes and the potential impacts to systems.

The digital identity models of three countries in particular – Estonia, Austria and the UK – embody the mix of regulatory, technical and commercial factors necessary to ensure a digital identity scheme conductive to the protection of privacy. In addition, two further schemes –

India's Aadhaar and Peru's eID scheme – are often held out as exemplary models. In section 10 we document these models and interrogate the factors that make them beneficial.

# APPENDIX C GLOSSARY

## C.1 Glossary of key terms

| Term | Definition |
|---|---|
| eID | Smart card based electronic identity |
| DPD | Data Protection Direction |
| FIDO | Fast Identity Online – the FIDO Alliance is an industry body building standards for authentication of individuals using mobile technology |
| OECD | Organisation for Economic Co-operation and Development |
| OAuth | Standardised protocol for sharing authorisation tokens |
| OpenID Connect | Standardised protocol for federated identity built on OAuth |
| PKI | Public Key Infrastructure |
| SIM | Subscriber Identity Module |

*END OF DOCUMENT*