Digital Identity Toolkit

Section 6: Data privacy and security

May 2023

What is this Toolkit?

Digital identity is a relatively new but rapidly evolving sector that can and will affect many aspects of our everyday lives.

Digital identities verify and authenticate someone's identity. They can then be used to access a wide range of services and opportunities, from health and education services, voting and travelling, through to online shopping and dating. Governments and the private sector are developing and implementing digital identity solutions, and they're likely to become increasingly common in the future.

While there is already a lot of information on this topic, much of it is in lengthy, technical reports and hasn't been collated into a simple format that non-technical people can understand. We hope this Toolkit can help close that gap.

This Toolkit has been designed to help you find everything you need to know about digital identity. Before producing it, we spoke with individuals and non-profits around the world to get a sense of what they'd like to know about digital identities.

The audience for this Toolkit is members of the public, non-profits, entrepreneurs, developers, journalists and academics who want to learn more about digital identity and how digital identities might be relevant to them in their lives or work.

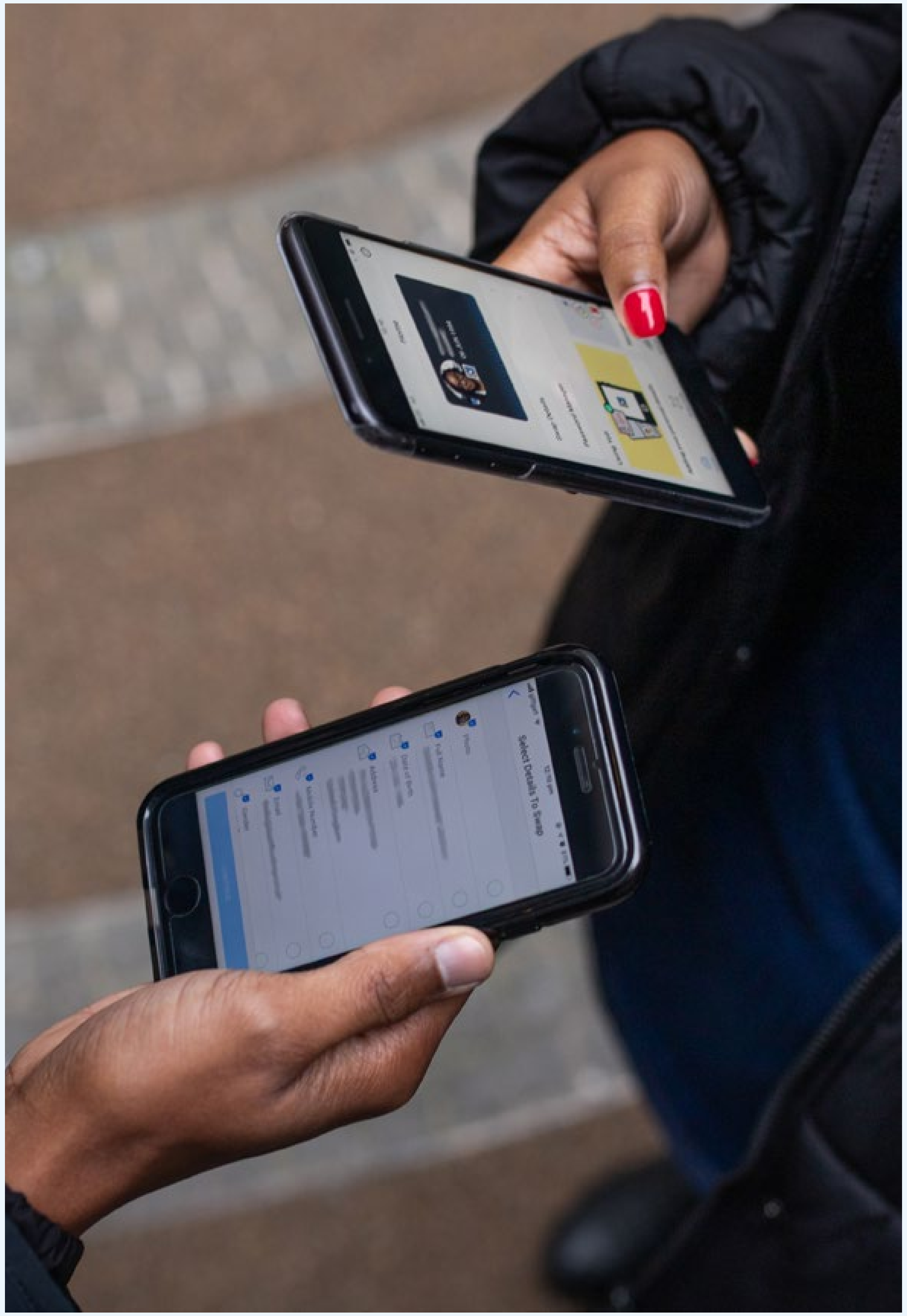
We hope you find this Toolkit helpful and welcome your feedback about how it could be improved.

Contents

Introduction

While digital identity can take many forms and be developed in many ways, it always involves people's personal information. It's therefore critical to consider both privacy and security matters when developing a digital identity system.

This chapter looks at the key privacy and security points that you need to factor in as you plan, develop and build the system. Your approach should involve determining what you are trying to achieve, and then looking at the most secure and privacy-friendly way to get there. This section has a detailed checklist of questions to help you protect your users' privacy and security at every stage, and to help you avoid falling foul of any data protection or privacy regulations.



Why privacy and security matters

In some jurisdictions, there are laws or regulations on privacy, data protection, biometrics, identity and children. Any digital identity system in these jurisdictions must comply with these rules.

As well as legal compliance, and in jurisdictions without such rules, a successful digital identity system needs to work as intended and be trustworthy. Anything that fails to consider relevant privacy and security matters is unlikely to be effective or trusted.

What follows is a useful checklist of questions to help you decide your reasons for using digital identity, the kind of information you will need and how you can protect your users' privacy and security at every stage through a review and testing process.

Before you start

The questions below will help you consider the different privacy and security aspects of your digital identity system. Answering them, assessing risks, considering mitigating measures and documenting the process not only helps to develop a robust formula, but will also provide evidence that you have properly assessed privacy and security matters and addressed any risks.

Purpose

What is it trying to do?

All identity systems should have a clear aim - you'll have to determine what problem you are trying to solve and whether a digital identity is the right solution. Ultimately, your solution needs to be justifiable, proportionate and suitable for actually solving the issue(s) at hand.

Are you providing identities or checking them? Do you need to identify people or authenticate them?

Setting up a person's digital identity requires more personal details than if you are simply authenticating them, and there are different requirements again if you are checking whether an identity is valid or not. If your system intends to do both or even all of them, you'll need to look at how you can separate the functions so that each works as intended and doesn't involve unnecessary data collection, data use, data sharing or data retention.

Do you need to report on any aspect of the system? If so, what information do you need to report, and to whom?

Part of its purpose may be to report on the population, monitor the take-up of IDs, permit access services and so on. You'll have to carefully consider what information is needed and why, and how to achieve that in the most privacy-friendly and secure way.

Data minimisation

What is the minimum information you need for each aspect of the solution?

One of the main criticisms of digital identity solutions is how much information is collected. Making sure that you only request, use, share or keep the minimum information necessary for each aspect not only protects individuals' privacy but also reduces the security risk. The larger or more varied the data collected is, the more attractive it is to hackers. Therefore, if this data is compromised, the worse the impact on the individuals affected will be. If you don't collect the data in the first place, there is less that can be potentially lost.

Another security and risk mitigation measure is to reduce the ability for someone else to identify an individual or connect different information about them. Consider how far you can remove identifiers, aggregate data (for reporting for example) or keep data in separate pots.

Are there any human rights or dignity reasons to minimise data or linkages?

The information gathered about people and their use of an identity system can provide a detailed picture of a person and their life. It could reveal sensitive or confidential information, such as using an ID at an HIV clinic.

Given the purpose of the system, its intended audience, use, who has access to the data, and what reporting you have to do, consider whether there will be any risks to a person's rights or dignity, and how you can mitigate that. This consideration upfront will also help inform and educate people about the solution and allow you to provide reassurances to those who have concerns about acquiring and using a digital identity.

Individual choices

What rights and choices do people have?

In some countries, there are privacy laws that provide certain rights with regard to people's personal information. If your system is intended to be used in multiple jurisdictions, consider whether to only offer rights where legally required or whether to offer the same rights to all. It is good practice to offer the latter - doing so will increase the trustworthiness and credibility of your system.

If there are no legally required rights in the jurisdiction, consider what rights to provide people. For example, it is sensible to allow people to check that the information you have about them is correct, as incorrect information can significantly affect decisions or entitlements to services. If there is incorrect information, there needs to be a way to correct it.

Users may also need to update information as their circumstances change - your solution should consider how they can do this. If it includes biometric information, consider how to update or overwrite this if needed. An example of this being a necessary feature is if a person suffers an accident that changes their appearance and needs to replace their biometric template.

It should also be built with an ability to delete information as needed. Some records may need to be kept for different periods of time, but in circumstances outside of your control, such as when a person dies, the system should allow for information to be deleted.

What is mandatory and what is optional?

This is related to data minimisation through deciding what information people must provide and what is voluntary. The design of the system should make this distinction clear and not prevent users from moving on to the next step in creating their digital identity if they choose not to give information that is optional.

How will users exercise those rights and choices? How can they update details or delete them?

Consider whether and to what extent individuals can check or update information themselves and how your system will authenticate them so that they can view or change their own records only. Also consider how you will verify updated information so that false information is not entered.

If people need to visit a specific location or contact a specific person or team, reflect on the logistics and practicalities of how they can do this, especially in remote communities or where there is a lack of infrastructure or transport. This prevents those in certain scenarios being unable to exercise their rights and choices.

Access control

Who needs access to what information and why?

The more people that have access to personal details, the more at risk those details are. For each aspect of your system, think about what people or roles need access to the data collected, and for what reason. Put in place a process for requesting, approving, enabling and disabling access, as well as a way to review access controls to check they are still needed and up to date. Review which roles may need permission to edit personal details, and which only need permission to view personal details, and to what extent these permissions are required.

How can you prevent unauthorised access?

Consider how those with the correct permissions will actually access the system, and what kind of identity check or login they will need. It is good practice to require two-factor authentication, and it is crucial to put in place a process to assign and revoke credentials.

Security

Where will you store the actual data? How will you keep it secure?

Think about how you will securely manage logins or add and delete users, administrators and other parties. You will also need to look into whether you will have local storage and backups or if you'll need to use cloud infrastructure. If so, will this be a public or private cloud? It is also necessary to consider what physical, technical and operational controls you will need. Information security frameworks such as ISO 27001 provide a useful checklist of appropriate measures.

Third parties

Is anyone else involved in developing and managing the system?

There could be many third parties involved to build and develop the technology, host the data, provide certain aspects, report information to, and manage enrolment. Identifying all of the parties involved will help keep you in control of the project and make sure you have addressed relevant privacy and security matters with each party.

Have you done enough diligence checks on the third parties?

The nature of the diligence required depends on what service the third party is providing for you. You will likely have varying requirements for third parties who will host or have access to personal details compared to one who may, for example, be providing you with cables and switches. You will have to prepare different sets of diligence questions that include matters such as privacy and security matters or perhaps more traditional concerns such as reliability (depending on the service provided).

Is it clear who has what role and responsibilities? Do you have agreements in writing?

Having this in writing allows you to set out the obligations and responsibilities of each party and provides you with some kind of redress if things go wrong. To protect you and those whose personal details are in the system, consider including these points in an agreement (where the other party hosts, receives or has access to personal details)

- The other party will not do anything with the personal details other than what you have asked them to do.
- The other party will tell you if there is a security breach or other incident that could compromise the personal details.
- The other party will make sure that their staff who are involved in the project understand that the personal details are confidential.
- The other party will have appropriate security measures in place (you may want to set out the minimum standards you require).
- If the other party needs to subcontract any work, they will carry out appropriate diligence on their subcontractors, tell you about them, and make sure they follow the same obligations and responsibilities.

Measure success

Do you know how you will measure success?

This is linked to the earlier point about reporting on aspects of the system. Once you have decided what is required, decide how best to collate and present the relevant statistics.

What data will you need to do this and how far can you de-identify, pseudonymise, aggregate or segregate data?

Linked to data minimisation, review what underlying data is necessary for your metrics and how you can present the right data to the right audiences in a way that does not risk identifying specific individuals.

Applicable laws

Are there any existing or draft laws, regulations, regulator guidance, codes of practice or other guidance that will apply to what you are doing?

These may cover privacy, personal information collection and use, biometrics, children's data, identity numbers, identity documents, security, encryption and more. Some countries also have references to privacy in their constitutions. Draft rules that are not yet concluded will provide useful information on the standards or practices that you will eventually have to comply with. You can future-proof your system and avoid expensive and time-consuming changes in the future if you design it with these upcoming rules in mind.

Is what you have planned in line with those requirements?

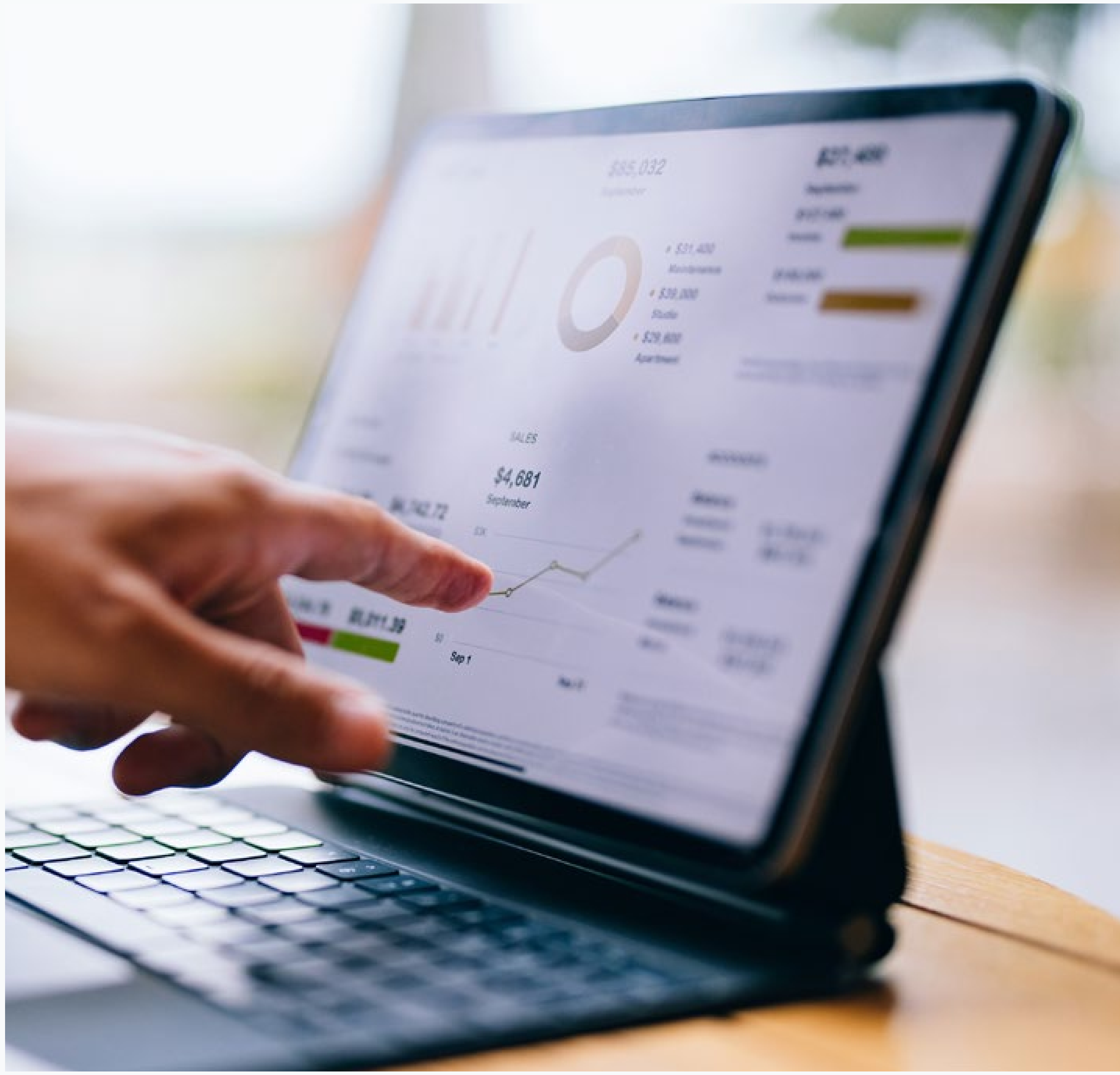
There are different ways to assess what you are planning against any of the requirements. Answering this list of questions and documenting your thinking, decisions, risks and mitigating measures is called a privacy risk assessment or privacy impact assessment. It is here where you'll be able to include information on any relevant rules and check whether or how you intend to meet them. Consider whether you need external help and what reliable resources may be available online. If the jurisdiction has a privacy regulator, they can often be a useful source of advice and guidance. Even where there is no regulator, there may be one in another similar or neighbouring jurisdiction whose online resources you could benefit from.

As you build

Changes to the plan

What needs to change and why? How does this affect privacy or security matters?

Most projects involve a degree of change as requirements are modified or added to, or as technology, time, funding and resources dictate what is and isn't possible. Therefore, it is useful to put in place a process to document and assess changes for privacy or security impact, and what steps you will take as a result.



Before you go live

Testing

Have you checked that everything works as intended?

A trial run helps to check for technology bugs or unintended consequences and ensures that the system makes sense to those who will use or run it.

Have you tested the security measures?

Testing that they work as intended avoids data issues and the subsequent problems when you launch. Your system will lose trust and credibility if it launches and immediately has a security issue.

Have you checked for data leakage, such as personal information in logs?

The metadata or data generated automatically by the use of a system, such as logs, can contain personal details if not set up correctly. Before you launch, check what is being collected so you have time to make any necessary changes.

Governanace

Have you got any required governance policies and processes in place?

Consider what you need for those who will run or administer the system as well as the individuals who will use it. Relevant policies and processes for administrators could relate to access controls, security, training, enrolment, reporting, dealing with individuals, exercising rights and choices and so on. For individuals, relevant policies and processes could cover enrolment, where and how to use the digital identity, or rights and choices.

For all parties involved, you will need to consider how they can ask questions, ask for help or make complaints.

Transparency

How will you tell people how the system works?

The trust and credibility of your system will depend on your efforts to inform and explain what it is, what it's for, how it works and why it's safe. Some jurisdictions may have legal requirements around what you have to tell people before you start collecting, using, sharing or keeping their personal details. Even if they don't, it is good practice to explain these aspects upfront in order to avoid fear, suspicion, mistrust, confusion or a refusal to engage.

Consider what communication strategies you will need. You may need multiple methods to inform people, such as in writing, verbally, or visually through signs, images or cartoons. Large-scale publication methods such as communicating through newspapers, the radio or the television could be options. Determine what is appropriate for the communications infrastructure in place and how you can avoid leaving certain people behind.

Also think about what you want to communicate. It may be enough to describe the system, its purpose and the data collection involved, but you may want to cover the benefits of taking part, how you are protecting their privacy and the security of their personal details. You may need to explain the technology or address specific cultural differences or concerns.

Do they know what to do, what not to do, where to find help and how to escalate matters?

Plan your training before you launch the system so everyone is prepared and confident of their role.

Review

How will you ensure that you effectively review your solution after it has gone live?

Decide when to review implementation and success after you launch. Make sure that you review and evaluate the system, its operation, security measures and the ability for users to exercise their rights and choices. It is crucial to develop an action plan for remediation and improvements and it is good practice to have this documented. Be sure to repeat these processes regularly to ensure that privacy and security rights are continually addressed and adhered to.

Case studies

Simprints

This is a very good example of a privacy-friendly and secure biometric digital identity. Simprints use fingerprints to provide a digital identity to individuals in the developing world. Simprints' mission is "to transform the way the world fights poverty. We build technology to radically increase transparency and effectiveness in global development, making sure that every vaccine, every dollar, every public good reaches the people who need them most".

https://www.simprints.com/

Estonia

Estonia has become known globally for its leadership and efforts to build a digital society of which identity is one part.

https://e-estonia.com/

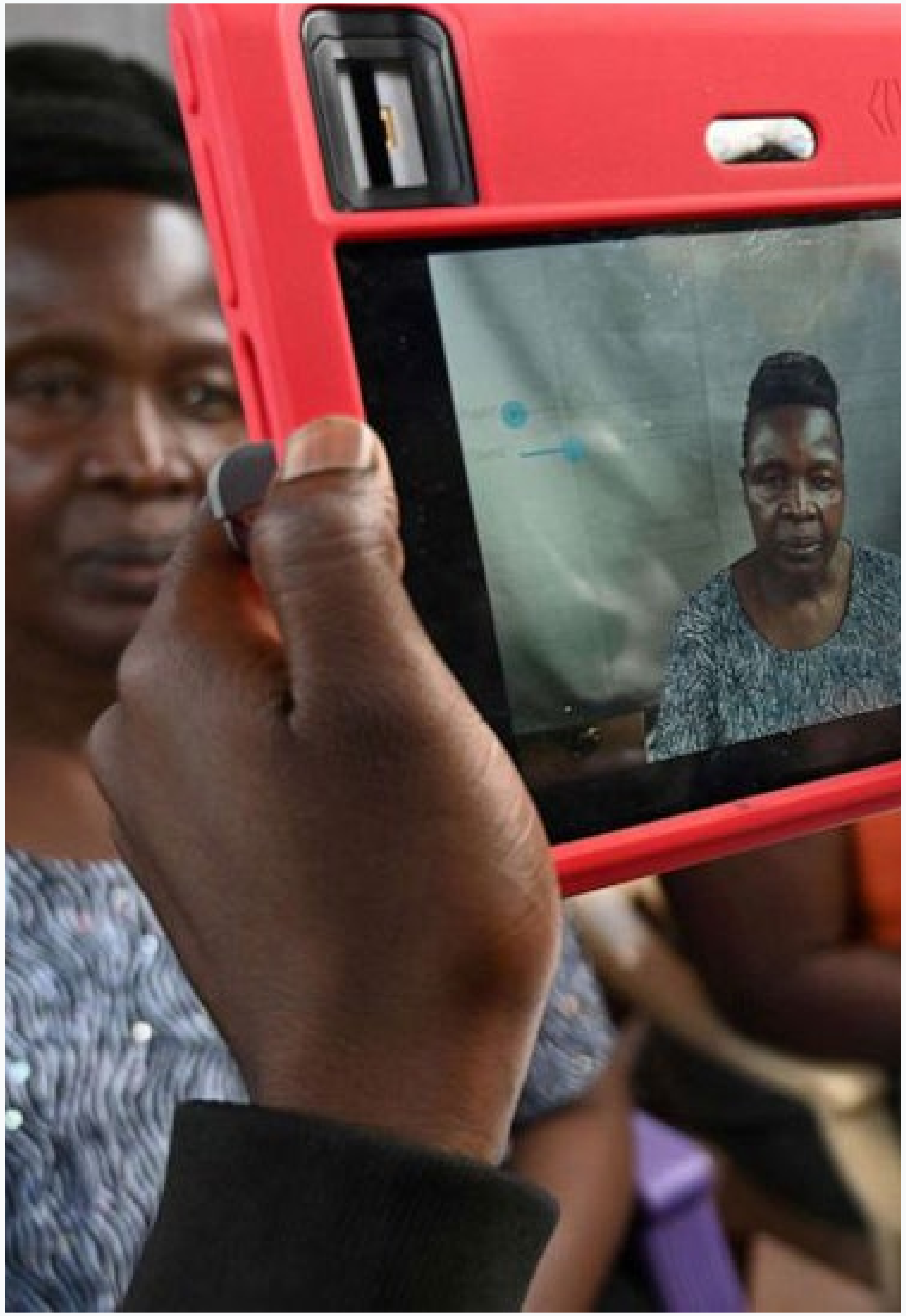
Kenya

Kenya's national identity system has been criticised for failing to consider privacy and security.

Kenya recently introduced the National Integrated Identity Management System (NIIMS) to create a central master population database. Individuals have to provide personal details including DNA, GPS coordinates of their residential address, retina scans, iris pattern, voice waves and earlobe geometry to get an identification document. NIIMS consolidate information in other government agency databases and generate a unique identification number known as Huduma Namba. This number allows citizens access to all government services, and the aim is to connect all parts of the country.

There are reports of misinformation and misunderstanding among the public, and most of the criticism has been of its invasiveness and surveillance potential. Other criticisms relate to the lack of a privacy law to protect the data and individuals' privacy, excessive data collection, security concerns and wide data sharing among the government.

The Kenyan High Court, while affirming the legitimacy of the voluntary registration scheme, prohibited the state from collecting DNA and GPS coordinates during the registration process.



Glossary

Term	Meaning
AML	**Anti-money laundering** checks are carried out by regulated businesses to perform due diligence and prevent financial crime.
API	**Application Programming Interface** refers to the software that allows for communication between two computer programs, such as applications, e.g. when Yoti shares your age with an app.
Back-end system	The infrastructure and system behind the 'front-end' of the digital identity solution. **API** would be a part of back-end system design.
Biometrics	**Biometrics** relate to the physical characteristics that can be used to identify individuals. Examples include fingerprint mapping, facial recognition or iris scans.
Blockchain	A way of recording information, so that it is stored across several computers connected in a network. This makes it almost impossible to exploit the system, creating a secure technology.
Cloud Infrastructure	The collection of elements needed for cloud computing. It includes hardware, software, network resources, computing power and storage.
GDPR	**General Data Protection Regulation** is legislation set out by the EU to protect the personal information of all data subjects within the region.

Term	Meaning
IDSP	**Identity Service Providers**, sometimes referred to as identity providers, allow people to remotely verify their identity.
KYC	**Know-Your-Customer** checks form a part of due diligence, which allow institutions to verify the identity of a customer whilst doing business with them.
MFA/V	**Multi-Factor Authentication/Verification** refers to a security measure in which the user must present at least two pieces of evidence to access a particular service. Alongside a username and password, the additional verification factor is usually based on one of the following things: something you know (e.g. a password), something you have (e.g. a mobile phone), or something you are (e.g. biometric data in the form of a fingerprint).
Open Source	This is a copyright licence under which the user can amend, use and distribute software. This is particularly helpful in easily creating digital identity platforms.
PII	**Personal Identifiable Information** is any data that can reveal someone's identity, either directly or indirectly. This must be protected at all times.
RP	A **Relying Party** refers to a server allowing access to secure software.
SDG	The UN has set out 17 **Sustainable Development Goals**. SDG 16.9 aims to provide legal identity for all, including birth registration.
SDK	A **Software Development Kit** is a collection of software development tools that makes it easier to develop an application, such as one for digital identity. It may also contain a software framework.

Further reading

Websites

- Data Protection Around the World (resource from CNIL, French regulator)
  https://www.cnil.fr/en/data-protection-around-the-world
- DLA Piper. Data Protection Laws of the World
  https://www.dlapiperdataprotection.com
- EU GDPR (privacy law)
  https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en
- European Privacy Regulators
  https://edpb.europa.eu/about-edpb/board/members_en
- Global Regulators
  https://pdpecho.com/the-list/
- **Information Integrity Solutions**
  An Australian privacy and security consultancy led by a former privacy commissioner that has published extensively on identity management.
  https://www.iispartners.com/identity-management
- Privacy Laws in Canada
  https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/02_05_d_15/
- **The Information Accountability Foundation**
  A global information policy think tank that works with regulatory authorities, policymakers, business leaders, civil society and other key stakeholders around the world to help frame and advance data protection law and practice through accountability-based information governance. They have published extensively on accountability and ethics frameworks.
  https://informationaccountability.org/
- **The Centre for Information Policy Leadership (CIPL)**
  A global privacy and security think tank based in Washington, DC, Brussels and London. CIPL works with industry leaders, regulatory authorities and policymakers to develop global solutions and best practices for privacy and responsible use of data to enable the modern information age.
  https://www.informationpolicycentre.com/
- UN Conference on Trade and Development. Data Protection and Privacy Legislation Worldwide
  https://unctad.org/en/Pages/DTL/STI_and_ICTs/ICT4D-Legislation/eCom-Data-Protection-Laws.aspx
- World Legal Information Institute. National Data Privacy Legislation
  http://www.worldlii.org/int/other/NDPrivLegis/

Reports

- *Baker Mackenzie Global Privacy Handbook*
  https://tmt.bakermckenzie.com/en/thought-leadership/global-privacy-handbook
- *Global Data Privacy Laws*
  https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3381593
- *International Conference of Data Protection and Privacy Commissioners*
  http://privacyconference2019.info/conference-report.pdf

Published by the Social Purpose Team at Yoti
www.yoti.com/social-purpose