Impact of GDPR on Identity and Access Management

Andrew Hindle (CIPP/E, CIPM, CIPT)

Table of Contents

Abstract

This article examines the implications of the General Data Protection Regulation ("GDPR", "Regulation") on Identity and Access Management ("IAM") process and system design. It introduces organisational and technical good practices that may help ensure demonstrable compliance with the Regulation as well as improve user experience and customer trust.

Although the focus here is on the GDPR, the approaches described may, by extension, also help in complying with data protection legislation in other geographies including (for example) the California Consumer Privacy Act ("CCPA"), or the Brazilian General Data Protection Law ("LGPD").

A Word to the Reader

We assume at least a basic knowledge of data protection and privacy - in particular, the GDPR[i] and the basic principles outlined in the OECD privacy guidelines[ii]. Even if you are not a security or privacy officer in your organization, understanding the rules will help you have better conversations with your privacy peers.

The privacy regulation landscape is evolving rapidly. Hence, the advice given here cannot be comprehensive and is neither intended nor should be considered as a substitute for legal advice. Whilst a good awareness of and sensitivity to privacy considerations is important for the digital identity professional, the majority of professionals are unlikely to be privacy lawyers. As with any area of regulation, it is always best to seek professional advice if at all uncertain.

Throughout the article, specific 'good technical practice' advice will be underlined; this same advice is also collated into a separate section at the end of the article as a checklist to follow for good IAM practices.

Introduction

Privacy conventions, regulations, and laws have been in existence for much longer than most people realise.[iii] As far back as 1948, the United Nations General Assembly enshrined a right to privacy in Article 12 of the Universal Declaration of Human Rights.[iv] In 1980, the OECD issued its "Guidelines on the Protection of Privacy and Transborder Flows of Personal Data,"[v] which were significantly revised and updated in 2013 (q.v.).

Individual countries and broader trading blocs continue to evolve their data protection and privacy regulations, in part to account for the evolution of technology. It is not uncommon for regulatory frameworks to lag behind technological developments, but changes will continue to be made in light of the impact that the Internet, connected devices, artificial intelligence, and genomics bring. To add to the complexity, the greater global mobility of individuals suggests that the local changes to data protection and privacy regulation impact changes in other jurisdictions. Aside from the changes in Europe, recent years have seen updated privacy regulations emerge in Brazil, Singapore, the Philippines, China, and parts of the United States, to name a few.

This evolution of regulation is important. When viewed through the lens of an ever-changing and evolving regulatory landscape, we can see the GDPR (and other modern privacy regulations) as a set of tools that can help us build better systems, not just as a set of checkboxes that we need to mark off.

Even if you work for an organisation that does not directly do business with Europe, certain elements of the GDPR have a global impact. Other privacy regulations may contain similar provisions; although it would be unwise today to plan for global harmonisation of these regulations, there are increasing commonalities between geographies. Recognize, too, that the GDPR applies equally to any data about individuals, whether it is data within a company about its employees or data about external individuals such as customers. In other words: the GDPR really does affect everyone.

The Regulation includes[vi] a Data Protection by Design requirement. Leaving aside the specific need to comply with the Regulation, these are fundamentally good design principles. They help mitigate business risk (e.g., the less data you have, the

less interesting you are to attack, and the less impact any attack will have), and they help reduce administrative overhead and wasted effort (e.g., the less data you have, the less likely it is that you will have duplication or contradictory records).

By definition, since the GDPR is concerned with personal data, these principles have significant implications for how we design and implement systems that use such data, including IAM systems and processes. Indeed, without an IAM foundation which itself complies with the Regulation, it's simply not possible for a final product to be compliant. (Though note that even if your IAM systems and processes themselves are Regulation-ready, you still need to ensure that your final service is compliant as well!)

The rest of this article will explore the principal considerations teams should have when developing IAM projects that can comply with the needs of the Regulation.

The Regulation applies to the physical representation of data (such as on paper) as much as it does to digital data. We'll focus here on digital information, but we'll make reference where appropriate to some specific implications (for example, in the areas of debugging, management reporting and so on.)

We'll start with some general observations, including commentary on your project team's composition and project structure. Then we'll focus on the four stages that data - including Personal Data - goes through during its lifecycle: create, read, update, and delete. For each of these stages, we'll reference some of the specific areas of the GDPR that apply and identify some architectures, tools, and techniques which can help. Where relevant, we'll note differences that might apply if you are a 'data controller' or a 'data processor' - but for the most part, the impact of these differences is more likely to be at the business/legal level, rather than the technical level. We'll finish with a summary of key takeaways that can also be used as a quick aide-memoire for future projects or team induction.

Terminology

- Data Mapping – "a system of **cataloguing what data you collect**, how it's used, where it's stored, and how it travels throughout your organization and beyond."[vii]

- Data Protection Officer ("DPO") – An individual who must be appointed in any organization that processes any data defined by the GDPR as sensitive.[viii] The DPO is responsible for "Working towards the compliance with all relevant data protection laws, monitoring specific processes, such as data protection impact assessments, increasing employee awareness for data protection and training them accordingly, as well as collaborating with the supervisory authorities."(See GDPR Articles 35, 37, 38, and 39 for more detail)

- Personal Data - Personal data are any information which are related to an identified or identifiable natural person.[ix] (See GDPR Article 4 (1) for more detail.)

- Data Protection by Design - data protection through appropriate technology and organizational measures.[x] See GDPR Article 25 for more detail.

General Observations

Collaboration with Privacy Advisors

With the principles of Data Protection by Design and Default, as defined by Article 25 of the GDPR, in mind, perhaps the most critical action you can take is to ensure that privacy requirements are considered at the very start of any project.

The GDPR requires many (but not all) organisations to have an appointed Data Protection Officer ("DPO"). Larger organisations may have a team of privacy advisors. If you're leading a project that involves data about individuals, it is your responsibility to make sure your privacy colleagues are involved. Make sure you involve the relevant people in your project at the very earliest stage. Remember that they may not know about your project unless you tell them! Even if you are not, don't be afraid to ask who is involved from a privacy standpoint, and then develop a working relationship with your advisor.

Your privacy specialist (if you do not have an experienced DPO) may not have a deep technical background, so you'll need to make sure you are providing them with the information they need in a format that makes sense to them so that they can provide complete and comprehensive advice. To make conversations more productive and efficient, consider doing additional privacy reading or even investigating publicly recognised qualifications or certifications from relevant privacy trade bodies or other institutions as part of your ongoing professional development.

Raising Concerns/Whistleblowing

If you are worried that your project or your organisation isn't taking privacy seriously enough, or if you think you've identified an issue that leaves you out of compliance with the GDPR, or - in the worst case - an actual data breach, make sure you know the right channels through which to report this. Larger organisations should have well-established reporting/escalation mechanisms and are also likely to have whistleblowing policies and processes which you can use as a last resort. Smaller organisations may not, and so you'll need to use your best professional judgement to work out how to most effectively raise concerns. **Do** keep good records of any such conversations but **do not** include specific examples of Personal Data when reporting issues if it can be avoided, lest you make a data breach worse (or unwittingly turn a potential incident into an actual data breach!).

Finally, if your organisation has a DPO, remember that the GDPR imposes quite strict requirements on the independent relationship and reporting lines of the DPO.[xi] These can be helpful reassurances if you find you need to escalate.

Personal Data – Definition and Mapping

The GDPR has a very broad definition of what is considered to be personal data: "'personal data'' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person."[xii]

It is important, then, to be equally broad in your approach. Make sure you fully understand across the entire project when data might be considered to fall into the category of personal data. Remember that even if you are dealing with aggregated or pseudonimised data, the Regulation will still apply if that data can be 're-identified'.

Consider a process known as "data mapping" at the start of a project to discover and map out what personal data might be used where and how and monitor and update this data map through the entire lifecycle of the project. Data mapping is **"**a system of **cataloguing what data you collect**, how it's used, where it's stored, and how it travels throughout your organization and beyond.**"**[xiii] Such a process is often part of a data protection impact assessment ("DPIA") but it can also be helpful in the overall design of your IAM architecture, so it's never wasted effort.

Individual tracking technologies

The use of individual tracking technologies including, but not limited to, cookies, is (at the time of writing) requires more than just the GDPR to consider if a service or website includes their use.[xiv] Each EU member state is responsible for issuing its own guidance in line with this directive, which has led to some important divergence in this area; case law is still evolving.

Some cookies can, however, fall under the GDPR - if, for example, they contain information which could be used (perhaps in conjunction with other data) to identify an individual. Hence, although perhaps not a core consideration for the IAM practitioner, it is worth being aware of and alert to potential issues in this area.

Special Circumstances

The guidance given in this article is intended to be generally applicable to all IAM projects. Some projects, however, will have particular circumstances which merit additional care and consideration. Some of these circumstances include children, law enforcement, and certain special category data. In all cases, practitioners should seek independent legal guidance.

Special Category and Other Sensitive Data

The GDPR makes special provision for certain more sensitive types of data, including (but not limited to) race, sexual orientation, political and religious affiliation, health related data and biometric data.[xv] If you are handling special category data in these areas, you will need additional safeguards.

Children

If your project involves data about Children, you will also need to take special care. The GDPR defines a Child as being below the age of 16 - but note that individual countries may (and some, including the UK, have) lower this limit to the age of 13 or below[xvi].

Law Enforcement and Personal Data

The use of personal data by law enforcement agencies is the subject of separate directives, regulations and laws; these are not considered in this article.

Automated Processing, Machine Learning, and Artificial Intelligence

The automated processing of personal data is a special circumstance in its own right, but one which merits a particular level of attention. Automated processing can include everything from machine learning ("ML"), algorithmic processing, the use of blockchain technologies, or artificial intelligence ("AI"). AI/ML, in particular, is a fast-moving field; developers, ethicists, lawmakers, and regulators worldwide and still trying to gauge the complete scope of what might be possible. It is already clear that AI/ML systems can infer or deduce 'facts' about individuals that were not part of the original data set (we often see this in user profiling). It's also clear that an original data set might not itself contain personal data (by definition) but can do once processed.

It's beyond the scope of this article to explore this area in any depth; the GDPR, however, imposes quite stringent requirements in the case of Automated Decision-Making and Profiling.[xvii]

Greenfield/Brownfield[xviii]

GDPR itself makes no distinction between greenfield applications and the refactoring of existing – 'Brownfield' – applications to bring them up to a state of compliance. From a practical standpoint, the former affords an easier path to using modern standards and techniques and is often less encumbered with legacy integration/support requirements.

Recognise, however, that GDPR compliance may drive you to review all the applications in your project's working environment. In some cases, there will be (more or less) simple technical or procedural solutions to achieve compliance. In others, however, you may need to revisit and revise the original business objectives in the light of the Regulation.

Proxy/Delegated Access

This article makes a general assumption that a data subject will be providing and/or accessing data about themselves. That said, there is naturally a variety of cases where someone might quite legitimately be accessing data on behalf of a third party.

In such circumstances, it is crucial to establish and apply appropriate mechanisms of authentication, identification, and authorisation (as recommended variously later in this article) both for the original data subject and for their proxy, along with delegation consent. In some circumstances, consent can arise via legal instruments such as a power of attorney, a court order, or similar. Establish whether this is a requirement for your use-case and design processes accordingly. You should also strongly consider maintaining a record of delegation consent and other authorisation actions where applicable. Standards such as UMA[xix] and Consent Receipt[xx] may help in this regard.

Backups

Having a reliable mechanism to secure your data in the event of a disaster is not only good general practice, it is also, essentially, required by the GDPR. Remember, though, that a poorly designed backup mechanism can potentially put you at greater risk of a breach. Ensure that data in any backup is protected with strong encryption and with other tools including, but not limited to, privileged access/user management – though be aware that these protections can complicate the restoration process. Certain sectors or applications may also require a physical or

paper-based backup mechanism. Whilst this is likely to be outside of the immediate scope of responsibility of the digital identity professional, do bear in mind that the GDPR requirements apply equally to data in physical form. Backups also introduce additional complexity in the area of retention.

Data Journey

Step One - Create

The first stage of our data journey - 'create' - starts at the moment you set out to collect personal data. Do not confuse this with the moment you write the data into a database (or other storage mechanism). Before you even request data from (or about) the data subject, you need to clearly understand and communicate to them what you are collecting and why, along with outlining their data subject rights. These are most commonly expressed via a privacy notice that uses clear and plain language - and you should at least ensure that the notice accurately reflects the way your system actually works!

Depending on the lawful basis for processing the relevant data, you may need to obtain the consent of the data subject for you to collect and process their information. How you obtain consent will differ from project to project, depending on what data is being collected and what it is being used for. Your privacy advisor can provide guidance.

From an audit perspective, consider keeping a record of that consent and/or providing your data subject with a record for themselves - evolving standards such as the Consent Receipt may be applicable here. Do remember, though, that any such receipt or record may itself contain Personal Data!

Create Minimally

If Data Protection by Design and Default formed the first guiding principle for your project, then your second guiding principle should be that of Data Minimisation. Data minimisation is good practice irrespective of compliance requirements: the less you collect or process, the less you have to protect and manage over time. It is also one of the 7 principles established by the GDPR for the handling of personal data.[xxi]

The bottom line: When collecting data from a data subject, collect and keep **as little data as you possibly can** in order to meet your requirements. Similarly, for

indirect data about a data subject, such as browser fingerprints,[xxii] collect and keep as little data as you possibly can. This means you need to have a good understanding of the business rationale for the project, so that you are clear about the justification and so that you can help your colleagues on the business side meet their obligations: it's always helpful to ask *why* a given piece of information needs to be collected.

Remember that the GDPR considers data in the aggregate. Consider whether there is any possibility of data your project is collecting being combined with other data the organisation holds in such a way as might result in identification of the individual (see also 'read' below). Avoid repeat collection of data that your organisation already holds about an individual. Aside from being a frustrating experience for the user, this also results in duplicate and/or conflicting records, which can cause problems with data accuracy, subject access requests, deletion, and other areas of the Regulation. If you have a large and disparate data map, consider using data discovery or meta-directory tools to help with visibility and consolidation.

Bear in mind that you may be collecting *implicit* or inferred data, which may also qualify as personal data: IP addresses, for example, or system analytics. These will need handling with the same diligence as data you *explicitly* request from or on behalf of a data subject. Even if this data is collected and used on a transient basis, it still needs handling correctly.

Consider also creative ways to limit the amount of data you collect. Besides simply collecting less, an organization might use an attribute service for answers to questions such as 'is the data subject over the age of 18', instead of collecting and storing the subject's date of birth, or requiring them to disclose credit card information. Technologies which can provide evidence that an authority has knowledge of certain information without revealing the information itself — zero-knowledge proof[xxiii], for instance — is also worth investigation. Be aware, however, that existing legal requirements may not yet take such technologies into account.

As noted earlier, this article mainly considers the impact of the GDPR on digital identity. However, the moment of data collection/creation is often where paper-based processes occur. Even if these are not your direct concern, it's no bad thing to make sure you understand how any paper records are being processed.

Possibilities for Federation

Having dealt with the basics, you now need to ask an important question: does your use-case actually need full user account creation. There is a tendency - born out of years of experience - to gravitate towards this as the first port of call in any identity project. Yet, in many cases, it's unnecessary; or it's something that only becomes needed later in the customer journey. Established standards like SAML or OpenID Connect support transient identity federation; this is often all you need. In such a case, you are only handling personal data (if at all) for a brief period of time, and so the normal data minimisation principles and precautions for data in transit may be sufficient. (use the most current version of TLS, plus additional specific data encryption as necessary)

If you do need a user account for technical reasons — session data persistence, for example — can it be made essentially 'impersonal' though the use of (for example) pseudonymous federation? Pseudonymisation allows for the user identity to be matched, using an identifier that cannot easily be associated with a known individual. Take care in this case, however: it can be possible to combine data in such a way as to re-identify the information, so defeating the purpose of pseudonymisation. Pseudononymous data is still considered personal data, and as such it must be considered against the requirements of the GDPR.

Storing Data

If you do find you need to persist data — whether pseudonymously or not — you will need to think about where and how you store the data. The usual protections for data at rest are important. Use appropriate encryption techniques and keep these under routine review: cryptography is an area of rapid development (particularly given the advent of quantum cryptography and the evolution of 'quantum-safe' algorithms and techniques). You should also ensure that the right processes are in place to keep supporting systems, applications, and libraries up to date and patched.

Other GDPR requirements notwithstanding, modern application design patterns will almost certainly lead you to provide an API for handling your personal data. In such cases, access to such APIs must be protected, ideally using a protocol such as OAuth; you could also consider using an API gateway. We'll come back to API protection again later in our data journey.

If you are considering a storage solution using a distributed ledger, you should take extra care. There is now clear consensus that storing personal data directly in such

a ledger is not good practice. Some solutions under development today may avoid this particular pitfall, but it is still worth bearing in mind, particularly if you are building your own. Until this area of technology is more stable, the best advice is to proceed with caution; to keep such projects under regular periodic review, even after deployment; and to ensure you have a well-documented and easily implemented way to reverse out of using the ledger-based solution, should that become necessary.

Using a cloud-based data or user store may have benefits from a risk management and privacy perspective. Ensure that you work with your privacy team so that your privacy notice accurately reflects the relationship between you and your provider.

Location of Data Storage

The GDPR does not itself impose requirements of data territoriality – that is, it does not require that data be stored in a particular geography – though regulations in other jurisdictions do. You should, at the very least, develop a flexible architecture that will allow you to segregate data on a regional basis should that become necessary — although bear in mind that this could mean collecting additional personal data which you might otherwise not need.

With that said, the GDPR **does** have requirements around the transfer of data outside of the European Union (i.e. to a "Third Country"). The transfer of personal data to any Third Country must always be a significant concern in the context of GDPR, and – although solutions can certainly be devised – this is an area of ongoing regulatory development. You will need careful discussion with your privacy adviser to make sure this is being handled correctly.

Step 2 – Read

Any and all access to the personal data you hold must be kept secure. At the most basic level, this means ensuring that you minimise any such access. If you are not already doing so, consider deploying a Privileged Access/User Management solution where applicable. You should also ensure that even those authorised privileged users, including database and systems administrators, cannot get access to personal data in clear form - even accidentally. Remember that **any** unauthorised access to personal data constitutes a potential data breach. Such a breach may be more or less severe and have greater or lesser consequences… but it is still a breach.

In order to provide useful functionality, whilst avoiding a potential data breach, be sure to use secure modern methods to authenticate and authorise your users, both internal and external. Use multiple factors of authentication; consider FIDO authenticators; avoid SMS as a factor; consider modern authorisation standards (and products which support them), including established protocols like XACML, newer standards like User-Managed Access ("UMA") and emergent approaches such as Transactional Authorization.

Note that 'authentication' is not necessarily the same as 'verification'. You may not need to establish the user's actual physical identity to any level of assurance in order to safely satisfy their request. However, where some level of assurance to a real-world identity is required, remember to treat any data used to verify the identity of the user with an appropriate level of security.

If you are pre-populating client-visible forms, be especially careful that such data is only displayed to the correctly authorised user, and that it cannot be cached across the visits of different users.

Modern application design patterns will likely mean that you have an API for 'read' operations. As noted earlier, any such API must be properly protected. Consider also adding additional program- or system-level protections: for example, protecting against multiple sequential reads by requiring additional authorisation or by imposing a total read limit or a repeat-time restriction.

Be conscious of other systems which may have access to personal data - security applications (especially ML or AI-driven solutions) and data mining tools, for example. Make sure such systems don't have unauthorised or unnecessary access to personal data in the clear, and be aware that in some cases, such access might constitute automated decision making or profiling (as referenced earlier).

Consider also unintended consequences. If you have a reporting tool which (for example) generates an Excel spreadsheet of data which can then be emailed, consider (a) whether all the PII needs to be in there; and (b) whether you can provide protection in some automated way upfront (for instance - by automatically creating an encrypted sheet, rather than relying on the user to have it do that for themselves), to help reduce the risk of an accidental breach further down the line.

Data Subject Access Request and Data Portability

The subject of the personal data has the right, under GDPR, to access the personal data you hold about them.[xxiv] This presents an obvious breach risk. If you are handling a response to a data subject access request, or if you are designing a system to be used for such a case, then you must be particularly careful to ensure that you correctly authenticate and/or verify the user; that they are properly authorized; and that the data you are sharing does not itself contain the personal data of other data subjects.

You are also required under GDPR to provide all the subject's personal data in a machine-readable format for data portability. The same security considerations apply in this case.

Somewhat perversely, in order to help satisfy some of these requirements, you may need to collect (or infer) more personal data than you might prefer, although you should always be careful not to collect more data that you absolutely need. For example: you may need to establish what country a given user lives in, is in, or is a citizen of, in order to establish what legislation applies! Depending on your system design, you can perhaps avoid storing this information and instead request it in real-time when the decision needs to be made (and verify it as needed).

Data Breach Reporting

Breach reporting is a special case in the context of 'read': if you are required to report a breach or a potential breach, you must ensure that you do not send personal data as part of your breach report. If you have automated breach or security reporting tools, make sure these tools don't accidentally create or worsen a breach by including personal data in their reporting. Consider also the use of privacy software solutions that can help search across data sets securely.

Step 3 - Update

GDPR mandates that data subject should be easily able to correct any personal data you hold about them. Make sure your system has such a mechanism. User self-service solutions can be particularly helpful in this regard, as long as they are appropriately easy to find and to use. Again, proper authentication and - in some cases - verification is crucial to mitigate against a potential accidental breach.

It is worth noting that this 'update' requirement of the Regulation may have implications for distributed ledger-based solutions. In particular, you should establish whether such a solution will allow for the rectification of a historical record in the ledger (or on the chain). Simply marking the historical record as 'no longer active' is unlikely to be sufficient.

Step 4 - Delete

In some instances, the GDPR provides the data subject with a right to request that the data you hold about them be deleted. You will need to make sure you have a straightforward way to do this - and that this mechanism is secured against accidental or deliberate misuse with appropriate safeguards including necessary levels and methods of authentication and authorisation. Consider maintaining audit logs for such transactions (bearing in mind that you will want to keep the actual personal data out of the log record), and potentially having a time-limited 'roll-back' mechanism in the event of an error.

The Regulation also requires that data be stored only for the period it is actually needed. Business requirements, informed by privacy needs, will dictate the length of the retention period; but you will need to design your system such that data can be easily expunged at the end of this period. Consider maintaining a separate record indicating when the data in question was originally created and running an automated task either to report on the data which has reached its retention date (hence flagging it for manual deletion) or to remove it directly.

For large and/or brownfield deployments, you may need to run a discovery process in order to establish what data you actually hold about a given data subject. There exists a variety of software solutions that can facilitate this.

As with 'Update', If you have an API (or other facility) which can perform data deletion - and especially if you allow for bulk delete - make sure you protect against misuse. For instance: add an additional (even a manual) check before a bulk delete or require additional authorisation for requests exceeding a certain number of records. You should also ensure you have a way to routinely back-up data and to restore in the event of a mistake (or a deliberate attempt to corrupt data), and consider forcing a backup via your API code before the delete process runs. Recall that retention of any such backup copies must be limited.

Conclusion

GDPR - and other modern data protection and privacy legislation and regulation - means we have to take extra care in designing, developing, and maintaining our IAM solutions. In particular:

- Collect only the data we need
- Only keep it for as long as it is needed
- Look after it when it is in our care
- Make sure it can only be accessed by those who should have access
- Make sure it can be appropriately updated
- Dispose of it safely when it is time to do so

We already have the tools we need to do this, but we need to be careful to apply those tools in the right way and to ensure that business owners aren't asking us to do things we shouldn't be doing:

- Only create accounts if absolutely necessary; use federation (SAML; OpenID Connect) or other transient or non-identifying information where we can (User Info; Zero-Knowledge Proofs)
- Authenticate users, preferably with strong and multiple factors of authentication (FIDO)
- Authorise users, preferably with modern protocols (XACML and UMA)
- Protect APIs (OAuth)

Much of what we need to do isn't new, and much of it has always been good practice. It's just not necessarily been standard practice or even top of the list for projects. New privacy regulations give us the opportunity to do things the right way.

Your IAM Project Checklist

- Ensure that privacy requirements are considered from the very start of a project, and routinely re-evaluated through the lifetime of the application
- Involve the relevant people (people who represent organizations consuming the IAM data as well as those serving as sources of truth for your IAM data, together with your privacy peers) in your project at the very earliest stage.
- Do keep good records of any conversations around potential data breaches but do not include specific examples of Personal Data when reporting issues.
- Map what, where, and how personal data might be used; this will be valuable input to a more complete Data Protection Impact Assessment (DPIA)
- If you are handling special category data as defined by GDPR and/or your local or sectoral privacy regulations, you will need additional safeguards.
- If your project involves data about Children, you will also need to take special care.
- Ensure that your organization's or service's privacy notice accurately reflects the way the system actually works!
- Collect the consent of the data subject for you to collect and process their information.
- Keep a record of that consent and/or providing your data subject with a record for themselves.
- Explore the Consent Receipt specification and emerging implementations.
- Collect as little data as you possibly can (data minimization).
- Avoid repeat collection of data.
- Consider using data discovery or meta-directory tools to help with visibility and consolidation.
- Explore zero-knowledge proof technologies and implementations and investigate whether such solutions should form a part of your deployment
- Instead of creating an account, consider instead using transient identity federation and/or single sign-on. If account creation cannot be avoided, consider using pseudonymous federation and/or single sign-on to reduce the amount of identifiable personal data you hold.
- Use the most current version of TLS plus additional specific data encryption as necessary.
- Use appropriate encryption techniques and keep these under routine review.
- Keep supporting systems, applications, and libraries up to date and patched.
- Protect access to APIs that handle personal data, ideally using a protocol such as OAuth.
- Storing personal data directly in a distributed ledger is not good practice.

- Develop a flexible architecture that will allow you to segregate data on a regional basis.
- The transfer of personal data to any Third Country (as defined in the Regulation) must always be a significant concern.
- Access (physical and digital) to the personal data you hold must be kept secure.
- Consider deploying a Privileged Access/User Management solution.
- Ensure that even those authorised privileged users, including database and systems administrators, cannot get access to personal data in clear form.
- Use multiple factors of authentication; consider FIDO authenticators; avoid SMS as a factor; consider modern authorisation standards (and products which support them), including established protocols like XACML, newer standards like UMA and emergent approaches such as Transactional Authorization.
- Be careful that Personal Data is only displayed to the correctly authorised user, and that it cannot be cached across the visits of different users.
- Be particularly careful to ensure that you correctly authenticate the user and that they are properly authorized.
- Avoid storing personally identifiable information, and instead request it in real-time when the decision needs to be made (and verify it as needed).
- If you discover a breach in your system, do not send personal data as part of your breach report.
- Make sure your system has a self-service mechanism to support the correction and/or deletion of a user's personal data.
- Consider maintaining audit logs for such transactions (bearing in mind that you will want to keep the actual personal data out of the log record).
- Consider maintaining a separate record indicating when the data in question was originally created and running an automated task either to report on data which has reached its retention date (hence flagging it for manual deletion) or to remove it directly, in line with your privacy policy and notice
- Check before a bulk delete and require additional authorisation for requests exceeding a certain number of records.
- Ensure you have a way to routinely back-up data and to restore in the event of a mistake (or a deliberate attempt to corrupt data), and consider forcing a backup via your API code before the delete process runs.

[i] For an overview, read the IDPro Body of Knowledge GDPR article. The full text of the regulation can be found at https://eur-lex.europa.eu/eli/reg/2016/679/oj

[ii] Organisation for Economic Co-operation and Development, "The OECD Privacy Framework," 2013, https://www.oecd.org/sti/ieconomy/oecd_privacy_framework.pdf

[iii] An overview of the history of the GDPR can be found at https://cloudprivacycheck.eu/latest-news/article/a-brief-history-of-data-protection-how-did-it-all-start/

[iv] United Nations, "The Universal Declaration of Human Rights," 1948, https://www.un.org/en/universal-declaration-human-rights/

[v] Organisation for Economic Co-operation and Development, "OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data," 2013, http://www.oecd.org/sti/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm

[vi] See Article 25 of the Regulation

[vii] KJ Dearie, "What is Data Mapping? The Importance of Data Mapping for GDPR Compliance," Termly, 30 October 2018, https://termly.io/resources/articles/gdpr-data-mapping/

[viii] "GDPR: Data Protection Officer," Intersoft Consulting, https://gdpr-info.eu/issues/data-protection-officer/

[ix] "GDPR: Personal Data," Intersoft Consulting, https://gdpr-info.eu/issues/personal-data/

[x] "GDPR: Privacy by Design," Intersoft Consulting, https://gdpr-info.eu/art-25-gdpr/

[xi] See in particular Article 38 of the Regulation.

[xii] Article 4 of the Regulation

[xiii] Dearly, "What is Data Mapping? The Importance of Data Mapping for GDPR Compliance," https://termly.io/resources/articles/gdpr-data-mapping/

[xiv] European Parliament, Council of the European Union, "Directive 2009/136/EC of the European Parliament and of the Council," November 2009, http://data.europa.eu/eli/dir/2009/136/oj

[xv] See Article 9 of the Regulation.

[xvi] See Article 8 of the Regulation.

[xvii] See Article 22 of the Regulation.

[xviii] Greenfield is a term used to describe a project with no prior work to constrain its development. Brownfield, in contrast, refers to projects with predetermined limitations based on having to work in an existing platform or under pre-existing constraints.

[xix] Specifications and Auxiliary Documents, User Managed Access Working Group, Kantara Initiative, https://kantarainitiative.org/confluence/display/uma/Specifications+and+Auxiliary+Documents

[xx] Lizar, Mark and David Turner, eds. "Consent Receipt Specification," Consent & Information Sharing Working Group, Kantara Initiativehttps://kantarainitiative.org/file-downloads/consent-receipt-specification-v1-1-0/

[xxi] "The Principles," Information Commissioner's Office Guide to the General Data Protection Regulation (GDPR), https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/.

[xxii] Katarzyna Szymielewicz, Bill Budington, "The GDPR and Browser Fingerprinting: How it Changes the Game for the Sneakiest Web Trackers," Electronic Frontier Foundation, 19 June 2018,

https://www.eff.org/deeplinks/2018/06/gdpr-and-browser-fingerprinting-how-it-changes-game-sneakiest-web-trackers

[xxiii] "Zero-knowledge proof," Wikipedia, last updated 24 January 2020, https://en.wikipedia.org/wiki/Zero-knowledge_proof.

[xxiv] "Art. 15 GDPR Right of access by the data subject," Intersoft Consulting, https://gdpr-info.eu/art-15-gdpr/