PRIVITAR
ENGINEERING PRIVACY

Getting ready for the GDPR

How data privacy software can help organizations take a fundamental approach to GDPR compliance

PRIVITAR
ENGINEERING PRIVACY

Introduction

The General Data Protection Regulation (GDPR) comes into force on the 25th May 2018. Research by the UK Direct Marketing Association indicates that 98% of companies believe the GDPR will affect them to some extent of which 44% expect to be very or extremely affected, yet 26% think their organization is extremely or somewhat unprepared. This points to a clear need for systems and expertise to help companies meet the challenging timeline for compliance.

Privitar offers enterprise data privacy solutions that equip organizations with the software they need to execute their privacy policies in consistent, scalable and transparent ways.

Privitar can help organizations achieve GDPR compliance, through:

1. Providing the tools to deliver on the fundamentals of privacy by design (PbD)
2. Direct compliance
3. Protection in the event of a data breach
4. Removing barriers to processing

Adopting data privacy principles not only helps meet compliance obligations, but also tackles the underlying drivers which the GDPR is itself responding to.

Data breaches are becoming more frequent and more expensive. Research by Risk Based Security shows that more files were exposed by data breaches in 2016 than the previous four years combined, and research by the Ponemon Institute shows that the cost of breaches has been increasing, from an average of $3m in 2013 to $4m in 2016.

Meanwhile consumers are becoming more privacy conscious, with research by TrustE showing 89% of consumers saying they avoid companies that do not protect their privacy[ii].

“98% of companies believe the GDPR will affect them to some extent, of which 44% expect to be very or extremely affected. Yet 26% think their organization is extremely or somewhat unprepared.”

[ii] Article 25 of the GDPR, Data Protection by Design and by Default, is an equivalent concept to PbD.

Privitar provides the solutions needed for organizations to deliver on Privacy by Design

The GDPR aims to provide the regulatory framework for EU data protection for many years (the 1995 Data Protection Directive it replaces will have lasted over 20 years). In that time both technologies and business models will change significantly. To ensure that the GDPR does not become outdated and a hindrance to innovation, the GDPR at times, focuses on fundamental rights and responsibilities, rather than prescriptive best practice. The resulting lack of clarity in some areas, for what compliance looks like in practice, is a challenge for some organizations.

Areas lacking detail should become clearer as guidance from supervisory authorities, best practice, codes of conduct, and certifications develop. These more agile instruments should bring clarity and a way for compliance requirements to keep up with rapidly changing business models and technologies. Until they are ready though, organizations will have to prepare with a level of uncertainty.

The best approach to dealing with this uncertainty is to look at the driving forces behind the GDPR: the need to mitigate privacy risks which have arisen out of the explosion in the use of personal data over recent decades. The GDPR forces organizations to respond to these risks. In areas where it is uncertain how this should be done, looking at the underlying risks and acting on these will deliver the outcomes intended by the GDPR, and is therefore likely to be the direction best practice and guidance goes in.

Privacy by Design is an approach that promotes the consideration of privacy protection from the start of a project. PbD is mandated by the GDPR. An important concept in PbD is data minimization, where privacy risk is mitigated by ensuring that only the data that is needed is processed. For many organizations this is challenging as they cannot easily separate what is useful in a dataset from what is personal.

The Privitar Data Privacy Platform provides a range of techniques for controlling privacy risks and data utility, so that projects do not carry more risk than is required. These capabilities range from pseudonymization to various types of anonymization and are designed to support transparency and auditability.

As these techniques mitigate the underlying risks, they respond to the driving concerns behind the GDPR, and are therefore one way of getting ahead.

Supports compliance directly

Anonymization

Recital 2[iii] of the GDPR states that data which has been anonymized to the point that it is no longer reasonably likely that an individual could be re-identified is no longer considered personal data, and is therefore outside of the remit of the GDPR.

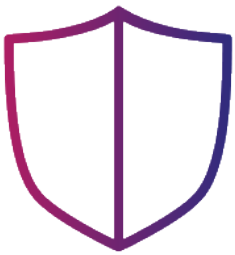
Privitar offers the software necessary to anonymize data in such a way that this criterion is met, thereby removing data from GDPR compliance requirements.

One way to achieve this is to create an anonymized copy of a dataset where direct identifiers have been replaced with pseudonyms and indirect identifiers have been blurred. The exact level of blurring required to meet the reasonable likelihood threshold is context specific and judgment is required. Privitar's expertise in data privacy helps organizations to find the right balance.

An additional challenge arises from different supervisory authorities potentially having differing views on what the appropriate level of anonymization is for a given example. In time the European Data Protection Board may issue general guidance ensuring a consistent approach for all authorities, which should address this.

[iii] In European law Recitals explain the reasons for an act and are used to help interpret how the Articles should be understood.

An alternative approach is to use differential privacy, a privacy guarantee currently being promote by Apple and Google. Privitar provides a solution that provides this guarantee by allowing controlled query access to datasets, rather than creating an anonymized copy. This approach is often used to provide less trusted parties with safe interactive access to sensitive data sets.



It is important to note that the processing required to anonymize data does not require consent. The Article 29 Working Group (A29WG) and the Information Commissioner's Officer (ICO) have both made it clear that the process of anonymization does not require consent. For further details, see p28 of the ICO's code of practice for anonymization.

Pseudonymization

Privitar's range of de-identification capabilities, includes pseudonymization which is explicitly called out as a key part of Data Protection by Design and Default (Article 25) as well as being a requirement for the security of processing (Article 32).

When it comes to pseudonymizing data, also known as data masking or tokenization, there are a number of factors to consider:

1. **Chain of custody.** Pseudonymization should be done within the data controller's environment. Sending the data to a third party to complete the pseudonymization before sending it back introduces new potential sources of risk.

2. **Consistent policy application.** It is important to apply pseudonymization policies uniformly rather than doing so on an ad hoc basis. Ad hoc pseudonymization can be inconsistent and leave data protection weaknesses. For instance, pseudonymizing one dataset does not effectively mitigate privacy risk if the same fields exist in another dataset being stored in the same environment which hasn't been pseudonymised. A policy based approach is therefore necessary to achieve, and demonstrate, consistency.

3. **Consistent pseudonyms to allow data set joins**. If a pseudonym is used for an identifier in one dataset, it is useful to have the capability to use the same pseudonym for that identifier if it is found in a different dataset. This ensures datasets which refer to the same objects can be linked after they have been pseudonymized, meaning joining and analysis on joined datasets can still take place.

4. **Policy based rather than tool set.** A clear policy-based interface which takes policies as inputs as opposed to commands, ensures that decision can be made by those who own the risk, understand the data, and know what they want to achieve from the processing. This isn't always possible with pseudonymization products which consist of a tool set usable only by technical staff.

5. **Scalability.** To be usable in today's big data world, pseudonymization solutions must be able to deal with the scale of large datasets, without significant performance issues.

6. **Reversibility.** While sometimes this is undesirable, in other instances it is essential. To maximize the number of instances where pseudonymization can be used to protect data, reversibility should be an option which can be toggled on or off as appropriate.

7. **Auditability.** To demonstrate to a supervisory authority that an organization has applied appropriate privacy policies, the application of those policies needs to be auditable. This can be achieved by tracking the lineage of datasets, recording which pseudonymization policies have been applied to which datasets and by whom.

Provides protection in the event something goes wrong

Fines and compensation

One of the reasons that the GDPR has received so much attention is the dramatic increase in potential administrative fines that the regulation brings in, with an upper limit of €20m or 4 percent of global annual revenue, whichever is higher. Article 83 outlines the considerations the relevant supervisory authority should take into account when deciding what size of fine is appropriate.

One of the factors that would be considered is the extent to which organizations have implemented the data protection requirements described in Articles 25 and 32, which explicitly call out pseudonymization, but would also include other, more advanced, data privacy protections.

Data de-identification reduces compliance risk in three ways. First through data minimization it can directly reduce the likelihood of incidents occurring. Second, it can reduce the size of fines by demonstrating responsible policies were in place. Third, it reduces the potential harm to the individual, reducing the motivation for both fines and compensation claims.

Data breach reporting (Article 34)

The GDPR introduces two new forms of data breach reporting. The first is reporting to the relevant supervisory authority, and the second to any individuals where the "*breach is likely to result in a high risk to the rights and freedoms of natural persons*". Notification of large numbers of individuals can be logistically challenging due to the scale involved, and harmful to the organization's brand due to the nature of the contact.

Individuals do not need to be notified if "the controller has implemented appropriate technical and organizational protection measures, and that those measures were applied to the personal data affected by the personal data breach".

The Privitar Platform allows organizations to create de-identified datasets that can prevent individuals from being identifiable in the event of a breach, and thereby reduces the potential impact on the individuals to whom the data relates. Privitar also provides audit trails to allow organizations to demonstrate that a policy has been applied to a given dataset. This can be strengthened by Watermarks, where Privitar inserts an indelible fingerprint in the data itself.

As well as being a deterrent to data being misused, Watermarks can be useful in the event of a breach. With even a small sample of the breached data, the Watermark can be used to identify what was in the original dataset, what privacy policy was applied to the data, and by whom.

Removes barriers to processing

Mitigating steps which support legitimate interests

A significant change made by the GDPR is to strengthen consent requirements. To avoid the challenges posed by these new requirements, it is expected that many organizations will shift from consent, where appropriate, to the 'legitimate interest' basis for processing. While processing on the basis of legitimate interest does not require consent, it does place certain other constraints on processing. The key criterion for legitimate interest processing is the balancing test.

This test weighs the legitimate interest of the organization against any potential risk of harm the processing may pose to the individual. If the result of the balancing test is that the processing presents too great a risk to the individual, then the organization controlling the data has the opportunity to try and rebalance the situation by taking actions to mitigate the identified risks to the individual (Article 6).

Data privacy software, such as Privitar, offers a way to mitigate risks to the individual arising from processing their personal data. So that, even if not executed to the level of full anonymization, organizations may be able to process personal data without consent, which otherwise would not be possible.

For more information on the balancing test, see page 42 of the A29WG's guidance.

DPIAs

Article 35 of the GDPR requires that prior to embarking on any new projects which may pose a risk to individual's rights, the data controller must first carry out an impact assessment to evaluate whether the processing is likely to pose a high risk to individuals. If they do find a high risk, then they must notify their supervisory authority of their plans prior to starting the processing (Article 36).

Much like the balancing test and the grounds for processing, data privacy software, such as Privitar, offers a way for data controllers to mitigate identified risks to the individual, and thereby potentially avoid needing to notify their supervisory authority or having to stop a planned project.

Right to be Forgotten

Article 17 requires that organizations delete the personal data they hold on an individual in certain situations. This can be undesirable for organizations, perhaps because it is difficult to actually delete the data, or because they have an interest in keeping the data.

In certain situations, Privitar can help by allowing organizations to delete the relationship between a pseudonym and an individual rather than all of the data associated with the pseudonym.

A fuller description of the Right to be Forgotten, and how Privitar can support compliance can be found in our whitepaper on the subject.

Third country data transfers

The GDPR keeps the rules on international data transfers much as they are (Articles 44-50). This means they continue to pose significant challenges for many organizations.

Using Privitar Lens, personal data held in one country can be queried interactively from another country, and differentially private aggregate responses returned, without any personal data being transferred across borders. This may offer a way for organizations to avoid the arduous requirements some have found with meeting Binding Corporate Rules or Standard Contractual Clauses.

Special categories of data

Certain categories of data have a special status in the GDPR. This includes data relating to children, data relating to a set of specified criteria, such as ethnicity or religion, or data relating to criminal convictions.

As there are different requirements for processing data of this kind, it may in certain instances be useful to drop these parts of a dataset. Privitar can be used to ensure that this redaction is done in a uniform and consistent manner.