Industry FAQs

International Association of Privacy Professionals (IAPP)
General Data Protection Regulation (GDPR)
Big Data Analytics
Webinar

Anonos Inc. (Anonos), in partnership with the International Association of Privacy Professionals (IAPP), hosted a webinar on enabling big data analytics under the GDPR that included industry experts representing regulatory, legal policy and technology perspectives:

Panelists:

Gwendal Le Grand
Director of Technology
and Innovation
at the CNIL

Mike Hintze
Partner at Hintze Law
Former Chief Privacy Counsel
and Assistant General Counsel,
Microsoft

Gary LaFever
CEO at Anonos
Former Partner at
Hogan Lovells

These industry experts respond to questions, which were not answered during the live event, submitted by the 600+ privacy professionals from around the globe who registered for the webinar. Following the webinar, interactions with companies and regulators revealed that most companies are at one of five stages of adjustment to the GDPR. Understanding these five stages can help companies that control and process personal data find a solution to their needs. These Industry FAQs include the following:

TABLE OF CONTENTS

SUMMARY OF GDPR MANDATED TECHNOLOGY CHANGES

- "Consent" and "Performance of Contract" alone will **not** provide legal basis for data analytics, artificial intelligence or machine learning using personal data under the GDPR.
- The GDPR moves beyond encouraging privacy by design to mandate **new requirements** for technical and organizational measures to protect fundamental rights of data subjects.
- Data assessments and inventories, along with reviews of policies and procedures **are** important but are **inadequate** by themselves to support new GDPR mandated requirements for data analytics, artificial intelligence and machine learning.
- Encryption, a critical data security measure strongly encouraged under the GDPR, will **not** support lawful analytics, artificial intelligence or machine learning using personal data.
- The GDPR provides a two year "grace period" that ends on May 25, 2018 to enable companies to develop and deploy **new** technology and organizational measures that support **"Data Protection by Default"** and "**Pseudonymity**" to control the linkability of data.
- What technical and organizational changes does your company have to make to do analytics, artificial intelligence and machine learning under the GDPR?

The GDPR represents a fundamental change in how data is processed. **Companies can no longer do what they did in the past with data linkages and expect to comply with the GDPR**. They must look at what steps they are taking to protect the rights of data subjects based on the uses of data they are making. Companies must have **new** protective mechanisms in place and show that they are giving controls to data subjects and that they are respecting data subjects' rights. **Compliance requires new technical measures**. Changes starting May 25, 2018 are as much about new technical requirements as they are about exponentially increased fines.

Three levels of de-identification discussed during the webinar are:

1. **Article 4(5) level de-identification** to support Data Protection by Default by satisfying GDPR requirements that additional information necessary to attribute pseudonymous data to data subjects is kept separate from the data and protected by technical and organizational measures;
2. **Article 11 level de-identification** to support Data Protection by Default requirements and excuse a data controller from data subject rights under Articles 15 to 20 by satisfying GDPR requirements that a data controller is not in a position to identify data subjects; and
3. **Anonymization level de-identification** to exclude data from GDPR jurisdiction by showing the inability to (a) single out, (b) link to, or (c) infer, the identity of data subjects. If these criteria are met, a data controller is on the "safe side." If these criteria are not met, a data controller must conduct risk analysis to prove that the risk of re-identification is sufficiently low; additional safeguards and techniques may be required.

The magnitude of new fines and joint and several **liability** among data controllers **and** data processors combined with the opportunity to embrace new technologies to improve business practices represents a tipping point that is not a negative – but a positive. So, while common data processing practices like using persistent identifiers that do not satisfy Article 4(5) pseudonymity requirements cannot be used as they have in the past, there are ways to continue business processes so that everyone can continue successful use of data.

QUESTIONS FROM PRIVACY PROFESSIONALS

Question No 1:

Several questions are grouped together for a combined response

1(a) "Gwendal may answer this, but does WP29 and the Commission understand that Controlled Linkable Data represents a new technical category for enabling data controllers and data processors to maximize data protection, utility and value in compliance with obligations under the GDPR? Are there efforts to get an opinion from WP29 (or EDPB)?"

1(b) "On page 14 of the Hintze/LaFever White Paper it says, "Documentation evidencing technical and organizational measures made possible by using Controlled Linkable Data are helpful in establishing the standard of care under the GDPR including … Article 40 - Establishing Codes of Conduct and Article 42 - Establishing Certification Mechanisms/Data Protection Seals/Marks." My questions are - when will codes of conduct and certificate mechanisms be established? How does my company confirm Controlled Linkable Data is well received by WP29? How can we hope to have it in place by May 2018 without guidance?"

Gwendal Le Grand

In 2014, WP29 published an opinion on anonymization techniques. This opinion does not go in that much detail with respect to the controlled linkable data. It is very unlikely that WP 29 or EDPB will draft an opinion on a specific product.

The GDPR will be applicable in May 2018. Certification is voluntary. A certification (by certification bodies or by the competent supervisory authority) is issued on the basis of criteria approved by the competent supervisory authority, or by the EDPB. WP29 is currently working on guidelines relating to certification.

A draft code of conduct can be prepared by associations or other bodies representing categories of controllers or processors. It shall be submitted to the competent authority, which shall provide an opinion and approve the code if it finds that it provides sufficient appropriate safeguards.

Gary LaFever

Interested parties should read the White Paper co-authored with Mike Hintze, **Meeting Upcoming GDPR Requirements While Maximizing the Full Value of Data Analytics** (available at anonos.com/whitepaper), which introduces Controlled Linkable Data as a new category of technical means for satisfying GDPR requirements.

Anonos BigPrivacy dynamic de-identification systems, methods and devices are examples of this new category called Controlled Linkable Data. This new category of data protection technology is also referred to as "Dynamic Data Obscurity" – for additional information, see http://informationaccountability.org/data-dynamic-obscurity-project/ and https://iapp.org/news/a/a-look-at-dynamic-data-obscurity/. Anonos licenses BigPrivacy dynamic de-identification systems, methods and devices to data controllers and data processors to help them maximize data protection, utility and value in compliance with GDPR obligations.

BigPrivacy dynamic de-identification systems, methods and devices help to satisfy the following requirements under EU data protection law:

- **Proportionality** under Article 52 of the EU Charter of Fundamental Rights;
- **Proportionality** under GDPR Recitals 4, 156 and 170 and Articles 6, 24 and 35;
- **Pseudonymisation** under GDPR Recitals 26, 28, 29, 75, 78, 85, 156 and Articles 4, 25, 32 and 89;
- **Data Protection by Default** under the GDPR Recitals 78 and 108 and Articles 25 and 47; and
- **State of the art** under the GDPR Recitals 78 and 83 and Articles 25 and 32.

Since 2012, Anonos has been actively engaged in research and development to advance the state of the art in data protection, privacy and security technology. Anonos BigPrivacy dynamic de-identification systems, methods and devices are protected by granted US patents (including, but not limited to, No. 9,631,481; 9,129,133; 9,087,216; and 9,087,215) and a portfolio of 50+ pending US and international patent applications. BigPrivacy dynamic de-identification systems, methods and devices technically enforce data protection policies using Digital Rights Management (DRM)-like technical controls down to the individual data element level. This is called Privacy Rights Management or PRM. Even in situations where data subjects are not directly involved, BigPrivacy dynamic de-identification technology manages risk to enable responsible use of data that respects the rights of data subjects.

Question No 2:

Several questions are grouped together for a combined response

2(a) "Based on my reading of the White Paper, persistent "anonymous" identifiers that we use at my company no longer legal under the GDPR. How are we supposed to do business?"

2(b) "My technology team uses persistent identifiers defined as Readily Linkable Data in the White Paper – what can I do to make them realize we have to change? How long will that change take?"

Mike Hintze

The GDPR does not automatically make persistent identifiers "illegal." But it does impose many new compliance obligations on all personal data. In effect, many of the requirements mean that data should be de-identified to the extent possible and consistent with the purposes for which is was collected. For some data uses, persistent identifiers are necessary. But for other uses, such as big data analysis, machine learning, etc., de-identified data typically can and should be used in order to meet GDPR requirements while preserving the utility of the data.

Gwendal Le Grand

GDPR just makes it explicit that some data that was sometimes called "anonymous" identifiers by some stakeholders are actually personal data and have to be protected as such. It means, inter alia, that the controller must have a legal basis to process them. This can be "legitimate interest" in such case, the controller will have to make sure that data subject can exercise their rights. It also means that these personal data must be protected with appropriate safeguards.

Gary LaFever

Companies can no longer rely on prior approaches or legal bases for data analytics, artificial intelligence, or machine learning. While consent remains a lawful basis under the GDPR, the definition of consent is significantly restricted. Recital 32 and Article 4(11) of the GDPR mandate that consent must now be "freely given, specific, informed and an unambiguous indication of the data subject's agreement to the processing of personal data relating to him or her."

These requirements for GDPR-compliant consent are not satisfied if there is ambiguity and uncertainty of data processing, as is the case with data analytics, artificial intelligence, or machine learning, or what I'll refer to as "big data." Heightened requirements for consent under the GDPR shift the risk from individual data subjects to data controllers and processors. Prior to the GDPR, risks associated with not fully comprehending broad grants of consent were borne by individual data subjects. Under the GDPR, broad consent no longer provides sufficient legal basis for big data. Data controllers and processors must now satisfy an alternate legal basis for big data processing.

In 2014, WP29 published an opinion on legitimate interests of data controllers which is still applicable under the GDPR. The 2014 WP29 opinion outlines the requirements for the legal bases for data processing found in Article 6 of the GDPR. If, as is the case with big data analytics, heightened legal requirements for consent under Article 6(1)(a) are not satisfied, then the requirements of one of the other legal bases must be satisfied for data processing to be lawful.

These other legal bases are:

- Performance of contract under Article 6(1)(b);
- Compliance with a legal obligation of a controller under Article 6(1)(c);
- Protection of a vital interest under Article 6(1)(d);
- Performance of a task in the public interest under Article 6(1)(e); and
- Legitimate interest under Article 6(1)(f).

The 2014 WP29 opinion clarifies that availability of performance of contract as a legal basis must be "interpreted strictly and does not cover situations where the processing is not genuinely **necessary** for the performance of a contract, but rather unilaterally imposed on the data subject by the controller." Scenarios in the WP29 opinion concerning limitations on permissible data processing in the context of a fictitious online pizza transaction clarify the limited availability of legal bases for data uses that are not genuinely **necessary** for a transaction.

The GDPR does provide a means to continue big data processing. If GDPR proportionality, necessity, and state of the art obligations are satisfied by complying with new Pseudonymisation and Data Protection by Default requirements, then legitimate interest under Article 6(1)(f) can be a valid legal basis for big data processing:

- **GDPR Article 4(5) defines Pseudonymisation as requiring separation of the information value of data from the means of linking the data to individuals.** The GDPR requires technical and organizational separation between data and the means of linking the data to individuals. Traditional approaches like persistent identifiers and data masking do not satisfy this requirement, since correlations between data elements are possible without requiring access to separately protected means of linking data to individuals. The ability to re-link data to individuals is referred to as the correlative effect, re-identification via linkage attacks. It is also referred to as the Mosaic Effect, because the same party who has access to data can link the data to individuals.
- **GDPR Article 25 imposes a new mandate for Data Protection by Default.** It requires that data must be protected by default and that steps are required to use it, as opposed to the pre-GDPR default, where data is available for use by default and steps are required to protect it. It requires that those steps enforce use of only that data necessary at any given time, for any given user, and only as required to support an authorized use, after which time the data is re-protected.

Part 2(b) of the question asks "My technology team uses persistent identifiers defined as Readily Linkable Data in the White Paper – what can I do to make them realize we have to change? How long will that change take?" In response to this question, I draw the reader's attention to Mike's answer to question 16 on page 17. In answering question 16, Mike responds "The 'grace period' is the two years between the passage of the GDPR in 2016 and May 2018. In May 2018, the GDPR will be fully enforceable and organizations should plan to be in full compliance."

The time for technology teams to transition from "Readily Linkable Data" to "Controlled Linkable Data" (as defined in the White Paper) is **now**. You can share with senior management that failure of technology teams to transition from Readily Linkable Data to Controlled Linkable Data exposes your company to "cease and desist" orders from data protection authorities that would force you to stop activities performed for decades imperiling the viability of ongoing operations. This is **in addition to** potential fines equal to the **greater of** 20 million Euro and 4% of global gross revenues, negative public opinion, class action law suits, etc.

Question No 3

"Privacy policies say, 'We do not store personally identifying information, we store only unique persistent identifiers that we use to track visitor and prospect activities.' If persistent identifiers are readily linkable back to identities of data subjects, how can they satisfy definitional requirements for pseudonymous data under Article 4(5) of the GDPR or Article 89 requirements for technical and organizational measures necessary to support secondary use of data like statistical analysis?"

Mike Hintze

The term "personally identifying information" is not used in the GDPR, and uses of such undefined terms have been criticized as being ambiguous or even misleading. The GDPR uses the term "personal data" and unique persistent identifiers of the type described in this question are personal data under the GDPR. Pseudonymisation under the GDPR means that the personal data in question "can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person." Thus, the definition of pseudonymous data does not preclude the presence of unique persistent identifiers. However, it is worth noting that some Pseudonymisation methods that maintain such identifiers may be considered relatively weak, and stronger methods that do not rely on such identifiers may further reduce risk and provide a stronger case for meeting the requisite technical and organizational measures.

Gwendal Le Grand

As Mike said, GDPR refers to the concept of "personal data" and pseudonymous data remains personal data. Article 89 clarifies that "Processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation. Those measures may include Pseudonymisation provided that those purposes can be fulfilled in that manner."

Gary LaFever

Companies often confuse the two concepts of "Personally Identifiable Information" or "PII," on the one hand, which is generally a US-based perspective focused on linked information and "Personal Data," on the other hand, as defined under EU law. Personal data under EU law (and the laws of other countries) extends beyond PII to also include linkable information which is a more expansive concept. Linkable information goes beyond data that by itself can be used to identify an individual (PII) to also include information that on its own may not be able to identify a person, but when combined with other information could identify, trace, or locate a person.

Privacy policies that protect **only** PII will **not** comply with GDPR requirements to protect the broader concept of personal data which includes both linked as well as linkable information.

Question No 4:

Several questions are grouped together for a combined response

4(a) "If I outsource my Data Protection by Default obligations to a data processor and they fail to comply with GDPR requirements, am I liable as the data controller or are they liable as the data processor responsible for that service?"

4(b) "Do I understand correctly that is a data controller relies on a cloud data processor for technology and that technology is not GDPR compliant for big data processing that they are both liable?"

Mike Hintze

Yes, under the GDPR both the data controller and the data processor can be liable for the non-compliance by the processor.

Gary LaFever

It is important to note that data controllers, joint data controllers and data processors have joint and several liability under the GDPR and can be liable for failures of one another under Recital 143 and Articles 26 and 82.

Question No 5:

Several questions are grouped together for a combined response

5(a) "In situations like indicated in the White Paper (employment-related data, IoT, etc.) how does a company prove legitimate interest to support secondary use of data/analytics?"

5(b) "If we cannot rely on consent alone, what do we have to do to prove legitimate interest to support secondary use of data? Is that where new GDPR requirements for organizational and technical measures come in?"

Mike Hintze

Under Article 6(1)(f) of the GDPR, determining whether "legitimate interests" can be a legal basis for processing involves a balance between the legitimate interests of the controller, and the fundamental rights and freedoms of the data subjects. It should be clear that the stronger the de-identification, the lower the risk to the data subject's fundamental rights and freedoms, thereby strengthening the case for a reliance on "legitimate interests." Further, Article 6(4) supports the idea that de-identification can be used to help justify a basis for lawful processing other than consent: "Where the processing for a purpose other than that for which the personal data have been collected is not based on the data subject's consent . . . the controller shall, in order to ascertain whether processing for another purpose is compatible with the purpose for which the personal data are initially collected, take into account, inter alia . . . (e) the existence of appropriate safeguards, which may include encryption or Pseudonymisation."

Gwendal Le Grand

The data controller should demonstrate its legitimate interest but most of all, it must prove that its interests are not "overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child." So, the data controller must balance this legitimate interest and the interest and fundamental rights of the data subject. Finding the right balance often requires providing tools for data-subjects to apply their rights (right to access, right to oppose…) and to inform the data-subject appropriately about such tools. Organizational and technical measures will also help finding the right balance.

Gary LaFever

Legitimate interest may be satisfied under Article 6(1)(f) if GDPR requirements for proportionality, necessity, and state of the art technology are satisfied by complying with Pseudonymisation and Data Protection by Default requirements. See answers to question 2 on page 5 for more details about the requirements for "legitimate interest" as a legal basis for processing data.

Question No 6

"What's the difference between anonymous data under the 95 directive and pseudonymous data under the GDPR?"

Mike Hintze

The meaning and treatment of "anonymous" data are essentially identical between the 1995 Directive and the GDPR. Under both, anonymity is a very high bar, meaning that the data no longer relates to an identifiable individual. Anonymization must be irreversible, with no known or foreseeable way to reconnect the data to an individual. Anonymous data is outside the scope of "personal data" and not subject to the 1995 Directive or the GDPR. What is new with the GDPR is an explicit recognition of intermediate levels of de-identification, including Pseudonymisation. Under the GDPR, "Pseudonymisation" means "the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person." It is a lower bar than anonymization, since Pseudonymisation typically can be reversed (although there are organizational and technical safeguards in place to prevent any unauthorized reversal). There are benefits to the use of Pseudonymisation under the GDPR, but pseudonymous data is still considered "personal data" and is still subject to the GDPR.

Gary LaFever

In response to this question, readers should review the summary of Gwendal Le Grand's presentation during the IAPP GDPR Big Data Analytics Webinar available at https://anonos.com/IAPP_Anonos_GDPR_Webinar_Summary_v1_3.pdf and copied below:

"The 2014 Article 29 Working Group guidance on anonymization (which is still relevant under the GDPR) acknowledges many benefits expected from big data but stresses that these benefits must be balanced against protecting the fundamental rights of data subjects which cannot be waived by data subjects. Advances in privacy friendly solutions and anonymization techniques are essential to ensure fair and effective competition and continued advances in permissible big data processing. Anonymization is a key trigger for big data because the rules for personal data protection do not apply to anonymous data which means that anonymization is an alternative to data erasure once the purpose of initial processing has been satisfied.

The 2014 Article 29 Working Group guidance on anonymization includes three criteria for assessing the efficacy of anonymization techniques – the inability to use the "anonymized" data set to (1) single out, (2) link to, or (3) infer, the identity of a data subject. If these three criteria are met, a data controller is on the "safe side." If these three criteria are not met, it does not mean that anonymization is not possible but a data controller must conduct a risk analysis to verify that the risk of re-identification is sufficiently low; additional safeguards and techniques may be required. Generalization and randomization techniques are means to help achieve anonymity. It is clear in the 2014 guidance – as it is clear in the GDPR – that pseudonymous data is not the same as anonymous data.

Pseudonymization (as well as encryption) are leading practices to ensure security of data (Article 32). Article 25 also recognizes pseudonymization as a means of achieving Data Protection by Default.

The GDPR provides additional triggers to facilitate desired economic uses of big data. The GDPR was designed with big data applications in mind as long as they respect the rights and liberties of data subjects. That is why there are categories of purposes (Articles 5-6) and permissible processing for compatible purposes as long as reasonable technical measures are in place to protect the rights of data subjects. Scientific, historical, and statistical purposes under Article 89(1) are considered compatible uses as long as technical measures are in place that ensure the protection of data subject rights and liberties. It is important to design technical measures by which data subjects can oppose specific uses and to provide transparency to data subjects about the ways their personal data will be used."

Question No 7

"Why can't preexisting technologies satisfy Data Protection be Default requirements?"

Mike Hintze

In some cases, they likely can. And different technologies may be appropriate in different scenarios. But the challenge for many organizations is to find technology tools that provide strong protections for data, are easily implemented, and preserve the value of the data.

Gary LaFever

For the first time, the GDPR specifically requires privacy by design under data protection legislation. However, Article 25 of the GDPR requires more than just privacy by design; it requires Data Protection by Default, **the most stringent implementation of privacy by design**. The GDPR further requires that Data Protection by Default be applied at the earliest opportunity (e.g., by pseudonymising data at the earliest opportunity) to limit data use to the minimum extent and time necessary to support each specific product or service authorized by an individual data subject.

A Data Protection by Default approach to de-identification leverages incentives built into the GDPR to use **new** technical measures to enable GDPR-compliant secondary use of data.

Traditional technologies like encryption, hashing, and Privacy Enhancing Techniques (PETs, e.g., k-anonymity, l-diversity, and differential privacy) were developed **long before** GDPR requirements were established. When used alone, encryption, hashing, and PETs fail to satisfy the GDPR's Data Protection by Default requirements. Static "anonymous/pseudonymous" tokens fall short because links between data subjects and identifying information remain readily ascertainable. Consequently, European regulators are likely to conclude that these techniques fail to satisfy Data Protection by Default requirements due to re-identification risks from linkage attacks and the Mosaic Effect. Finally, stateless "anonymous/pseudonymous" tokens developed for PCI compliance in the payment card industry fail to enforce re-linking and revealing of personal data under the controlled conditions necessary to support iterative analytics and secondary uses of data.

The GDPR provides incentives to use **new** technical measures that enable the flow, commercial use, and value maximization of data in a way that recognizes, respects, and enforces the rights of individuals. This may be accomplished by controlling the linkability of identifying information to individual data subjects = Controlled Linkable Data.

Question No 8:

Several questions are grouped together for a combined response

8(a) "Can I rely on encryption to satisfy technology requirements under the GDPR for using big data?"

8(b) "Under the GDPR, can we consider that data encryption (not database encryption) is a Pseudonymisation or Anonymisation technique to make safe personal data?"

Mike Hintze

Encryption is an important security tool, and as such can be a key part of meeting security obligations under the GDPR. But is different than Pseudonymisation and anonymization. Encrypted data may still be directly tied to an individual. By contrast, Pseudonymisation is designed to obscure the connection to the individual, and anonymization aims to irreversibly break the connection to the individual.

Gary LaFever

Encryption is an invaluable security tool, use of which is encouraged under the GDPR. However, **encryption alone does not satisfy obligations under the GDPR with regard to Data Protection by Default or Pseudonymisation**.

Question No 9

"The White Paper makes it look like Data Protection by Default or Controlled Linkable Data technology does more than GDPR compliance. If it helps open new areas of business, can I make an argument that budget should come from biz dev and other business functions?"

Gwendal Le Grand

Article 25 is about data protection by design and by default. Under article 25 (2) "the controller shall implement appropriate technical and organizational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed. That obligation applies to the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility. In particular, such measures shall ensure that by default personal data are not made accessible without the individual's intervention to an indefinite number of natural persons".

So, Data Protection by Default is actually required under GDPR.

Gary LaFever

Previously, penalties imposed for data practices that failed to respect the fundamental rights of EU data subjects were so minimal that some companies engaged in "regulatory arbitrage" and simply paid the fines rather than complying with EU data protection obligations. The magnitude of liability under the GDPR is amazingly large given administrative penalties as high as 4% of global gross revenues plus joint and several liability among data controllers **and** data processors. These fines and penalties are this high because EU legislators and regulators want data controllers and data processors to take the fundamental rights of data subjects seriously.

While discussions at your company may start with the strong economic downside of not complying, privacy professionals have the opportunity to present what GDPR changes can mean in a positive way which may enable you to get more engagement from management on a positive approach. **You need to show them that it is no longer discretionary and it is no longer optional**. Technologists need to be at the table together with privacy professionals, as well as people responsible for generating revenue and value through data. The good news is that Controlled Linkable Data enables a scalable solution that operationalizes privacy policies needed to comply with the GDPR and other global data protection regimes and enables innovative data uses that increase data value and utility. Through discussions with such a stakeholder group, you can have a productive discussion.

Question No 10

"Are privacy policies no longer important?"

Mike Hintze

Privacy policies continue to be important under the GDPR. In fact, compared to the 1995 Data Protection Directive, the GDPR imposes much more extensive obligations to provide specific information to data subjects. So, updating privacy statements / privacy policies to reflect these new requirements will be an important part of GDPR compliance.

Gary LaFever

The GDPR now requires technical and organisational measures to safeguard the rights and freedoms of EU data subjects. This does not mean that privacy policies are no longer important. Rather, it means that **policies by themselves** are no longer sufficient. New Data Protection by Default and Pseudonymity requirements under the GDPR impose new obligations to **technically enforce privacy policies**.

Question No 11

"Does differential privacy satisfy requirements of Data Protection by Default?"

Mike Hintze

The GDPR doesn't specify any particular technology, and depending on the circumstances, different technologies may be appropriate for meeting the requirements. Organizations should look at their particular data needs and determine what technologies can both demonstrate strong data protections while preserving the value of the data and enabling the purposes for which the data was collected.

Gary LaFever

The White Paper co-authored with Mike Hintze, **Meeting Upcoming GDPR Requirements While Maximizing the Full Value of Data Analytics** (available at anonos.com/whitepaper) provides an overview of the shortcomings of pre-existing technologies to satisfy GDPR requirements. The answer to question 7 on page 11 provides details regarding the inability of pre-existing technologies like differential privacy to support GDPR requirements for Pseudonymity and Data Protection by Default.

Question No 12

"At what step in the data collection phase should de-identification begin?"

Gwendal Le Grand

When you process personal data, you must have a retention period associated to it. Once this has expired, you must either erase or anonymize the data. If you don't have the right to collect the data, then anonymization does not help.

Pseudonymising/anonymizing the data should be done as soon as possible according to privacy by default/design.

Question No 13

"If Pseudonymisation is not anonymization, what is the value in using pseudonymous data?"

Mike Hintze

Unlike fully anonymous data, pseudonymized data is still considered "personal data" under the GDPR and therefore subject to its requirements. However, as described in the White Paper and in my paper on de-identification and the GDPR (available at https://ssrn.com/abstract=2909121), the use of Pseudonymisation can be an important part of meeting GDPR requirements. And the stronger the Pseudonymisation or de-identification, the greater the risk reduction and the more valuable it can be in meeting those obligations.

Gary LaFever

The GDPR represents a fundamental change in how data is processed. Companies can no longer do what they did in the past with data linkages and expect to comply with the GDPR. Personal data can no longer be used the same way. Companies must have new protective mechanisms in place and show that they are giving controls to data subjects and that they are respecting data subjects' rights. This requires new technical measures – Data Protection by Default did not exist prior to the GDPR. Pseudonymisation and de-identification are means to enjoy greater data usage rights under the GDPR. Three levels of de-identification discussed during the webinar are:

1. **Satisfy Data Protection by Default obligations** – Article 4(5) level de-identification supports Data Protection by Default by satisfying GDPR requirements that additional information necessary to attribute pseudonymous data to data subjects is kept separate from the data and protected by technical and organizational measures.
2. **Avoid rights of access, rectification, erasure, restricted processing, and data portability** – Article 11 level de-identification supports Data Protection by Default requirements and excuses a data controller from data subject rights under Articles 15 to 20 if a data controller can demonstrate it is not in a position to identify data subjects.
3. **Avoid GDPR jurisdiction** – Anonymization level de-identification excludes data from GDPR jurisdiction by showing the inability to (a) single out, (b) link to, or (c) infer, the identity of data subjects. If these criteria are met, a data controller is on the "safe side." If these criteria are not met, a data controller must conduct risk analysis to prove that the risk of re-identification is sufficiently low; additional safeguards and techniques may be required.

Question No 14

"If Big Data (including personal data) is obtained via a data broker, what are the obligations of the purchaser regarding such data?"

Mike Hintze

Any personal data, no matter the source, in is generally subject to the same GDPR obligations. It may be prudent for a purchaser to ensure that the data broker has provided assurances that it has made the data available in compliance with all applicable privacy and data protection laws, and has obtained all necessary consents from the individuals affected. Further, the notice requirements under Articles 14 and 15 of the GDPR specify that where personal data is obtained from a source other than the data subject, the data controller must provide notice of the categories of personal data obtained, and the source(s) from which the personal data originate, and if applicable, whether it came from publicly-accessible sources. Thus, purchasers of personal data may consider updating their privacy statements to include such information if necessary.

Gwendal Le Grand

Also, when processing is based on consent, "the controller shall be able to demonstrate that the data subject has consented to processing of his or her personal data" (article 7). The buyer should ask for the proof of this consent to demonstrate that consent has been obtained.

Gary LaFever

It is important to note that data controllers, joint data controllers and data processors have joint and several liability under the GDPR and can be liable for failures of one another under Recital 143 and Articles 26 and 82. If a company acquires data from a data broker and that data broker does not have a permissible legal basis to the data, **both the data broker and the data broker customer – as joint data controllers – may be liable under the GDPR for penalties, class action law suits, etc**.

Question No 15:

Several questions are grouped together for a combined response

15(a) "I would appreciate input on what corporate stakeholders should be invited to a GDPR big data compliance kick-off meeting?"

15(b) "It seems that GDPR compliance requires both a compliance as well as a technology budget. How do I know what people to get involved? Where do I start to assemble the right team?"

15(c) "Which executive roles should tackle the GDPR big data prep?"

15(d) "Do I need to get a "business owner" to take lead in shepherding this through my organization? Who would that be?"

Mike Hintze

Every organization is different, so it's hard to generalize on the best way to effectively fund, adopt, and implement compliance measures within an organization. But many organizations do find it helpful to have business leads, who can partner with legal and compliance professionals, to drive adoption of new policies, procedures, and/or compliance tools and technologies. And this partnership typically works best when it happens at all levels within the relevant teams – including at the executive level.

Gary LaFever

See answers to question 9 starting on page 12 for a discussion about the need to support Data Protection by Default and the benefits of assembling a broad-based stakeholder group who will appreciate the benefits of enabling a scalable solution that operationalizes privacy policies needed to comply with the GDPR and other global data protection regimes while enabling innovative data uses that increase data value and utility.

Question No 16:

Several questions are grouped together for a combined response

16(a) "Are we supposed to start compliance efforts on May 25, 2018 or already be fully compliant by that date? Is there a grace period?"

16(b) "Why do we have two years to comply with the GDPR? What is the expectation of where we will be on 25 May 2018?"

Mike Hintze

The "grace period" is the two years between the passage of the GDPR in 2016 and May 2018. In May 2018, the GDPR will be fully enforceable and organizations should plan to be in full compliance.

Question No 17

"The business tells me that the real value of the data we collect is because of the connection to individuals. When using this de-identification this value is gone. What is your take on this?"

Gwendal Le Grand

If you can demonstrate that there is a need for processing the non-de-identified data and **that you have a legal basis to do so**, you may not have to de-identify the data (see the response about legitimate interest).

Gary LaFever

This question incorrectly assumes that "connections" to individuals are only possible by using "Readily Linkable Data" or "Linked Data" as defined in the White Paper that Mike Hintze and I recently wrote **Meeting Upcoming GDPR Requirements While Maximizing the Full Value of Data Analytics** (available at anonos.com/whitepaper.) Readily Linkable Data and Linked Data as defined in the White Paper can no longer be used the same way they have in the past. You must have **new** protective mechanisms in place and show that you are giving controls to data subjects and that you are respecting their rights and this requires new technical measures – Data Protection by Default did not exist prior to the GDPR. As Gwendal Le Grand said during the webinar, what changes on May 25, 2018 is as much about new requirements as it is about fines. The combination of the magnitude of fines, potential liabilities and penalties together with the opportunity to embrace new technologies to improve new business practices hopefully represents a tipping point that is not a negative – but a positive. So, while things like persistent identifiers cannot be used as they have in the past, there are ways – like Controlled Linkable Data – to continue business processes so that everyone can be successful.

Question No 18

"We saw a note on consent for children in Mike's slides, could they speak to that in more detail if possible?"

Mike Hintze

The GDPR has added a parental consent requirement that is similar (but not identical) to what exists in the United States under the Children's Online Privacy Protection Act (COPPA). Specifically, under the GDPR, where the legal basis for processing personal data is the consent of the data subject, such consent is valid only of the data subject is at least 16 years old. If the data subject is under the age of 16, the consent must be given or authorized by a parent. EU member states can adopt a lower age – but no lower than 13. The data controller must make reasonable efforts to verify that consent is given or authorized by the parent, taking into consideration available technology. Companies may be able to leverage an existing parental consent process designed for COPPA, but there will likely need to be different implementations for European users, with different age triggers and possibly different means of verification.

GLOSSARY OF KEY TERMS

CONTROLLED LINKABLE DATA

[The following is from the White Paper entitled Meeting Upcoming GDPR Requirements While Maximizing the Full Value of Data Analytics]

- Controlled Linkable Data allows data use and the unlocking of data value in a way that enables compliance with the GDPR, all while enhancing individual data subject privacy. In brief, Controlled Linkable Data represents a potential cornerstone technological approach for data controllers, regulators and data subjects for three key reasons:
  1. **It decouples data elements from re-identifiable linkages to data subjects while enabling selected portions of that data (or abstracted ranges or groupings of that data) to be available to legitimate, authorized users for specific times or purposes – or in specific places** – in a way that can be fully GDPR compliant (this means providing technical and organizational measures to enable GDPR-compliant secondary uses of data like analytics, machine learning and artificial intelligence);
  2. **It defines, enables and manages multiple, different levels of de-identification, based on situation-specific applicable regulatory and policy implications, producing the highest practical level of de-identification for any given data use**, and optimizing the balance between maintaining the utility of data and protecting privacy and security; and
  3. It enables data controllers to meet the precise exclusionary standards of GDPR Articles 11(2) and 12(2), because the **de-identification performed uses unique technological means to sever the re-identification links between data elements and data subjects;** this renders data controllers without re-identification key access "not in a position to identify data subjects," thus meeting the exclusion criterion set forth by the GDPR.

DATA PROTECTION BY DEFAULT

- With reference to Data Protection by Default, GDPR Recital 78 stipulates that:
  - "The protection of the rights and freedoms of natural persons with regard to the processing of personal data require that appropriate technical and organisational measures be taken to ensure that the requirements of this Regulation are met. In order to be able to demonstrate compliance with this Regulation, the controller should adopt internal policies and implement measures which meet in particular the principles of data protection by design and Data Protection by Default. Such measures could consist, inter alia, of minimising the processing of personal data, pseudonymising personal data as soon as possible, transparency with regard to the functions and processing of personal data, enabling the data subject to monitor the data processing, enabling the controller to create and improve security features. When developing, designing, selecting and using applications, services and products that are based on the processing of personal data or process personal data to fulfil their task, producers of the products, services and applications should be encouraged to take into account the right to data protection when developing and designing such products, services and applications and, with due regard to the state of the art, to make sure that controllers and processors are able to fulfill their data protection obligations. The principles of data protection by design and by default should also be taken into consideration in the context of public tenders."

- With reference to Data Protection by Default, Article 25 of the GDPR provides that:

  1. Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures, such as pseudonymisation, which are designed to implement data-protection principles, such as data minimisation, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects.
  2. The controller shall implement appropriate technical and organisational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed. That obligation applies to the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility. In particular, such measures shall ensure that by default personal data are not made accessible without the individual's intervention to an indefinite number of natural persons.

PROPORTIONALITY

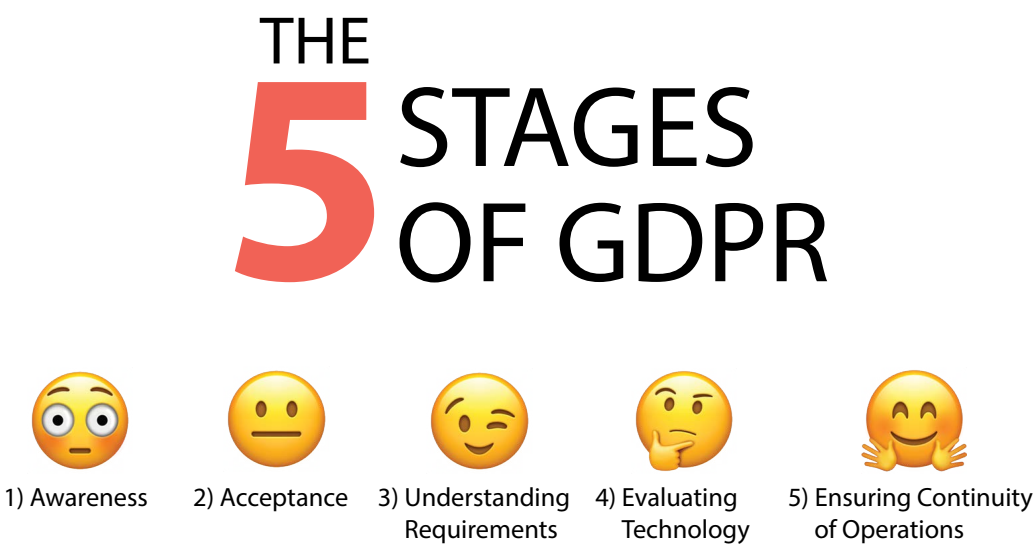
- Article 52(1) of the Charter of Fundamental Rights of the European Union (2010/C 83/02) provides that "Any limitation on the exercise of the rights and freedoms recognised by this Charter must be provided for by law and respect the essence of those rights and freedoms. **Subject to the principle of proportionality**, limitations may be made only if they are necessary and genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of others."
- GDPR references to "proportionality" include Recitals 4, 156 and 170 and Articles 6(4), 24 and 35.
- There are four generally recognized stages to proportionality under EU law, namely:

  1. There must be a legitimate aim for a measure;
  2. The measure must be suitable to achieve the aim;
  3. The measure must be necessary to achieve the aim, **that there cannot be any less onerous way of doing it**; and
  4. The measure must be reasonable, **considering the competing interests of different interested parties**.
     (See P Craig and G de Burca, EU Law (5th edn OUP 2011) 526)

PSEUDONYMITY

- With reference to Pseudonymity, GDPR Article 4(5) provides that:
  - "Pseudonymisation means the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person."

5 STAGES OF GDPR ADJUSTMENT



By: Gary LaFever, CEO at Anonos

This article includes new information responding to questions raised following publication of a prior version appearing in International Association of Privacy Professionals (IAPP) Privacy Perspectives.

On January 31, I took part in an IAPP-hosted webinar on managing risk and big data analytics under the EU's General Data Protection Regulation alongside Gwendal Le Grand, the Director of Technology and Innovation at the CNIL, France's data protection authority, and Mike Hintze, former Microsoft chief privacy counsel and now partner at Hintze Law. The IAPP GDPR Big Data Analytics Webinar was entitled "How to Comply with the GDPR While Unlocking the Value of Big Data."

Based on interactions with companies and regulators following the webinar, Anonos found that companies are at varying stages of adjustment to the upcoming regulation. Understanding these five stages, we believe, can help companies that control and process personal data find a solution to their needs. So, here we go:

Stage One – Awareness

At this stage, a company is aware that the GDPR contains new protections for EU data subjects and threatens significant fines and penalties for non-compliant data controllers and processors. Since compliance enforcement commences in the spring of 2018, many companies are stuck at this stage and are postponing moving to a solution. Yet preparation for compliance with the GDPR should begin now.

The company is usually also aware of the GDPR's broad jurisdiction. It applies to all companies processing personal data for one or more EU citizens, regardless of where the company is located or has operations. The company is also aware that penalties for noncompliance can include fines of up to four percent of global gross revenues, along with class-action lawsuits, and direct liability for both data controllers and data processors for data breaches, data breach notification obligations, and so forth.

Stage Two – Acceptance

A company at this stage realizes it cannot rely on prior approaches or legal bases for data analytics, artificial intelligence, or machine learning. While consent remains a lawful basis under the GDPR, the definition of consent is significantly restricted. Under the GDPR, consent must now be "freely given, specific, informed and an unambiguous indication of the data subject's agreement to the processing of personal data relating to him or her."

These requirements for GDPR-compliant consent are not satisfied if there is ambiguity and uncertainty of data processing, as is the case with data analytics, artificial intelligence, or machine learning, or what we'll refer to as big data henceforth. These heightened requirements for consent under the GDPR shift the risk from individual data subjects to data controllers and processors.

Prior to the GDPR, risks associated with not fully comprehending broad grants of consent were borne by individual data subjects. Under the GDPR, broad consent no longer provides sufficient legal basis for big data. Data controllers and processors must now satisfy an alternate legal basis for big data processing.

Stage Three – Understanding Requirements

At this stage, a company appreciates that the GDPR does provide a means to continue big data processing. If GDPR proportionality, necessity, and state of the art obligations are satisfied by complying with new Pseudonymisation and Data Protection by Default requirements, "legitimate interest" can be used as a legal basis for big data processing:

- GDPR Article 4(5) defines Pseudonymisation as requiring separation of the information value of data from the means of linking the data to individuals. **The GDPR requires technical and organizational separation between data and the means of linking the data to individuals. Traditional approaches like persistent identifiers and data masking do not satisfy this requirement, since correlations between data elements are possible without requiring access to separately protected means of linking data to individuals.** The ability to re-link data to individuals is referred to as the correlative effect, re-identification via linkage attacks. It is also referred to as the Mosaic Effect, because the same party who has access to data can link the data to individuals.
- GDPR Article 25 imposes a new mandate for Data Protection by Default. **It requires that data must be protected by default and that steps are required to use it**, as opposed to the pre-GDPR default, where data is available for use by default and steps are required to protect it. **It requires that those steps enforce use of only that data necessary at any given time, for any given user, and only as required to support an authorized use, after which time the data is re-protected.**

Stage Four – Evaluating Technology

A company at this stage is evaluating technology to determine if it satisfies GDPR requirements for both Pseudonymisation and Data Protection by Default.

- Pseudonymisation requires separating the information value of data from the ability to attribute the data back to individuals.
- Data Protection by Default requires revealing only that data necessary at a given time, for a given purpose, for a given user, and then re-protecting the data.

The Advent of Controlled Linkable Data

The IAPP GDPR Big Data Analytics Webinar introduces the term Controlled Linkable Data to describe **new** technology that achieves two critical goals:

- **Supports Pseudonymisation by separating the information value of data from the ability to attribute the data back to individuals; and**
- **Satisfies Data Protection by Default, defined as revealing only that data necessary at a given time, for a given purpose, for a given use, and then re-protecting the data.**

How does Controlled Linkable Data achieve these dual goals? It replaces restricted data elements, such as personal data under the GDPR, protected health information under HIPAA, and contractually restricted elements by dynamically changing pseudonymous tokens. As a result, Controlled Linkable Data does not enable correlations or "linkage attacks" back to the identity of individuals without access to keys.

Controlled Linkable Data also provides access to more accurate data, because it unlinks it without degrading it. On the other hand, alternative approaches apply Privacy Enhancing Techniques (PETs) indiscriminately, without allowing for what purpose the data will be used for, a shortcoming that degrades the value of the data.

Controlled Linkable Data can be implemented with a revolutionary approach, called Anonos BigPrivacy. Five granted patents, plus 50+ patent applications worldwide protect the unique capability of Anonos technology to use dynamically changing identifiers to enable Controlled Linkable Data.

Stage Five – Ensuring Continuity of Operations

Companies at this stage of adjustment are seeking to verify that technology vendors satisfy GDPR requirements for Pseudonymisation and Data Protection by Default, so that by using the technology they can ensure ongoing continuity of operations.

At what stage are you? Let us know. Contact us at: LearnMore@anonos.com

---

The Anonos intellectual property program, which is backed by 5 granted patents and 50+ additional patent applications worldwide, is the only technology that assures ongoing access to dynamically changing identifiers. This assurance produces invaluable benefits:

- Companies do not have to worry about disruptions in access to Anonos' BigPrivacy implementation of Controlled Linkable Data due to claims by third parties.
- Companies also have the option of using Anonos' BigPrivacy technology the way that best suits their needs.

Note:

If a vendor other than Anonos supports Pseudonymisation and Data Protection by Default, which are the two requirements for Controlled Linkable Data, by means of dynamically changing identifiers, the interested company should verify the vendor has rights under Anonos patents.

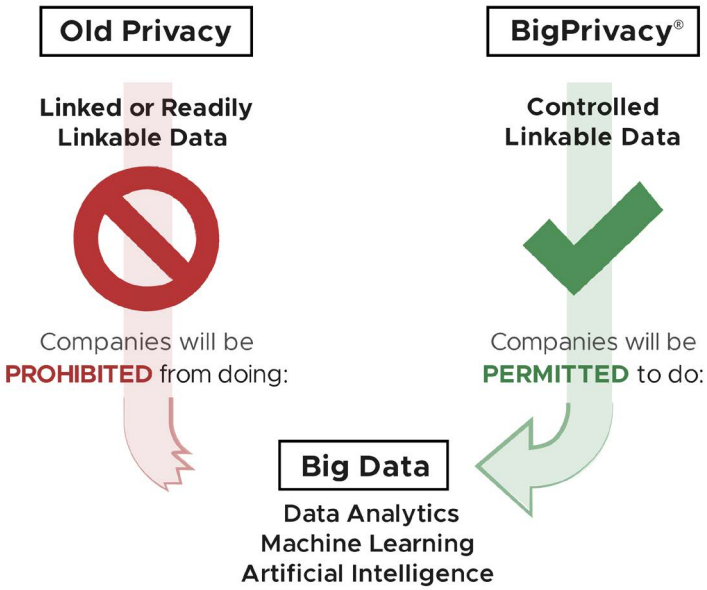
If a vendor supports Pseudonymisation and Data Protection by Default by means other than dynamically changing identifiers, the company expressing interest should verify that this alternate approach satisfies the requirements for proportionality and necessity under GDPR Recitals 4, 156, 170 and Articles 6(4), 24 and 35, as well as state of the art requirements under GDPR Recitals 78 and 83 and Articles 25 and 32.

GDPR locks up the value of Big Data.

BigPrivacy® unlocks it.

Since all major companies rely on Big Data for analytics, machine learning, and artificial intelligence, it's critical that companies now start to enable compliance in a way that protects and increases the irreplaceable value of their big data assets.

Unlocking the Value of Big Data



BigPrivacy is a patented technology that protects privacy by automatically "Anonosizing" data and actually expands the amount of data available for responsible use.

Click to learn more about unlocking the value of Big Data analytics under the GDPR

Or email us at:
LearnMore@anonos.com