IAPP GDPR Big Data Analytics Webinar

Summary

January 31, 2017

The GDPR represents a fundamental change in how data is processed. Companies can no longer do what they did in the past 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. Personal data can no longer be used the same way. Companies must have 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. Changes starting May 25, 2018 are as much about new technical requirements as they are about increased fines.

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

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 can demonstrate it 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 be successful.

This webinar is for informational purposes only and is not intended to, nor shall it be construed as, providing any legal opinion or conclusion; does not constitute legal advice; and is not a substitute for obtaining professional legal counsel from a qualified attorney on your specific matter. Opinions expressed in this webinar do not represent official positions of La Commission Nationale de l'Informatique et des Libertés (CNIL) or Hintze Law. All trademarks, service marks, trade names, trade dress, product names and logos in this document are the property of their respective owners.

A. Panelist Presentations

Presentation by Gary LaFever, CEO at Anonos and former Partner at Hogan Lovells

The GDPR requires fundamental changes in how personal data is processed. However, if done correctly, the GDPR enables positive new business opportunities as well. Data protection by default requirements permeate the GDPR and expand upon traditional notions of data minimization and minimum necessary data. Subject to fines up to 4% of global revenue, implementation of technical and organizational mechanisms are required to ensure that only the necessary personal data required for each specific processing purpose – whether collection, scope of use, length of storage, or accessibility – are used. The White Paper co-authored with Mike Hintze, Meeting Upcoming GDPR Requirements While Maximizing the Full Value of Data Analytics (available at https://anonos.com/whitepaper), introduces a new technical means of satisfying "data protection by default" requirements under the GDPR – Controlled Linkable Data.

Controlled Linkable Data enables data to be used for a range of purposes while preserving privacy and protecting data from unauthorized processing. Controlled Linkable Data facilitates extraction of the full value of data, enabling both GDPR and other regulatory compliance as well broad data utilization. Companies will benefit from proactively implementing Controlled Linkable Data to satisfy GDPR data protection by default requirements while maximizing data value and minimizing compliance risk and liability.

Existing technologies were not developed to satisfy GDPR requirements for data protection by default. Traditional approaches involving Linkable or Readily Linkable Data – like the use of persistent identifiers – will prevent companies from doing big data analytics, machine learning or artificial intelligence under the GDPR. Controlled Linkable Data, on the other hand, enables big data analytics, machine learning and artificial intelligence to be performed.

Presentation by Mike Hintze, Partner at Hintze Law and former Chief Privacy Counsel and Assistant General Counsel at Microsoft

De-identifying data serves as both a means of complying with the GDPR and maximizing the value of data. The GDPR goes significantly farther than the 95 EU Data Directive by imposing new obligations on data controllers and processors - both in terms of substantive obligations (meeting new data subject rights) and procedural obligations (data protection by default, data protection impact assessments, data breach obligations). The GDPR makes it harder to rely on data subject consent by itself and puts more focus on satisfying GDPR requirements for legitimate interest to process personal data for big data analytics, machine learning or artificial intelligence.

The magnitude of GDPR penalties (up to 4% of global gross revenues and joint and several liability between data controllers and data processors) make compliance an economic imperative.

GDPR compliance requires adoption of new technology but there are clear paths forward to address requirements under the GDPR, including use of pseudonymization, de-identification and anonymization to help comply with these requirements. The GDPR recognizes a full spectrum of different gradations that exist around de-identification. This is a positive step beyond the binary approach of the 95 EU Data Directive where data was either personal data or anonymous data. Under the GDPR, there are gradations between those two extremes that are recognized. Personal data still includes identified data (i.e., data that is clearly tied to the identity of a data subject) as well as identifiable data (i.e., data that includes at least a theoretical means of re-linking to the identity of a data subject).

The GDPR recognizes two different concepts of identifiable (vs. identified) data. The first is the explicit inclusion of "pseudonymous data" which is defined under article 4(5) as requiring processing of data in a manner that can no longer be attributed to an individual data subject without the use of additional information provided such additional information is kept separately and subject to technical and organizational measures that ensure it cannot be attributed to an individual. An even stronger level of de-identification is recognized in Article 11 when a data controller can demonstrate it is not able to re-identify a data subject. The GDPR also recognizes the highest level of de-identification – anonymous data – in which case GDPR restrictions do not apply.

Adopting de-identification as a compliance mechanism has numerous benefits under the GDPR because the GDPR incorporates a risk-based approach to data protection. The stronger the level of de-identification applied to data, the lower the risk of re-identification. You can use de-identification to help demonstrate adoption of data protection by default. If you meet the higher-level Article 11 type of de-identification, you get relief from data subject rights under Articles 15 - 22. De-identification also helps meet security obligations that are risk-based. In addition, de-identification provides a stronger case for relying on legitimate interest as a basis for processing data (as opposed to data subject consent by itself) since it helps to balance interests of a data controller and interests of data subjects.

Presentation by Gwendal Le Grand, Director of Technology and Innovation at the CNIL (French Data Protection Authority)

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.

B. Questions & Answers Discussed Live During Webinar

Question No 1

Three questions were read together for a combined response

1(a) Most businesses are focusing only on bare minimum "tick in the box" exercises rather than using this as an opportunity to transform the way they manage and use personal data. What would be your advice to them?

1(b) How do I make our technologists understand why we must process data the day the GDPR goes into effect differently than the way our company has processed data for years prior to the GDPR?

1(c) My company's technologists use the lack of specified requirements and specifications under the GDPR as an excuse not to change what we do and how we do it – any suggestions?

Gary LaFever

This is an important set of questions and this is a wake-up call. The GDPR is not a rule that enables you to make minor changes. It requires a fundamental shift - hopefully slide 7 from the webinar presentation deck helps to convey this. It does require data protection by default which has never been required before. Previously, penalties have been so minimal that companies engaged in "regulatory arbitrage" and simply paid the fines versus complying with data protection obligations. The magnitude of liability under the GDPR could be 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. Discussions with management at your company should start with the strong economic downside of not complying. This requires that privacy professionals on this webinar act as the standard bearers who say "this isn't something that we can simply take a 'tick in the box' approach to." Therefore, slide 4 was added to the webinar presentation deck - if you present what GDPR changes can mean in a positive way you will get more engagement from management on a positive approach. This will likely require changes to architecture - technologists hate that. 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, together with people responsible for generating revenue and value through data. Through discussions with such a stakeholder group, you can have a productive discussion.

Mike Hintze

I agree with what Gary said. There is a temptation and a natural reaction in many cases to say "we have not been handed a clear roadmap by the regulators that we have to take steps 1, 2 and 3 – it's all very amorphous so we are just going to throw up our hands and do nothing." That's exactly the wrong thing to do. While it seems that May 2018 is a long way out but it isn't given the types of things that need to be done to get ready for the GDPR. A lot of companies are just getting started and others are trying to wrap their heads around what this all means but companies need to start doing something. They need to start putting steps in place. When the regulator comes "knocking" after May 2018, a company must be able to say "these are the things that we did." At the end of the day there is going to be some uncertainty and things are going to have to be "tweaked" as more guidance comes out. As the GDPR starts to be enforced we will all collectively have a better understanding of what it is going to look like in practice. But taking steps now to deal with new process requirements, steps to deal with how data is stored, managed and processed - those are going to be important steps that must be started now. You can't just "flip a switch" and do these things over night. You can't just wake up in April of 2018 and say "oh, now it is time to get GDPR compliant." It is going to take some time.

Gwendal Le Grand

The GDPR is applicable in 2018 as you said. That means we have 15 months left which is not a lot of time because there are a couple of things that must be changed in your organization regarding the governance of privacy. I would say that it is clear that companies must get prepared. There are also new rights for individuals - this is not the topic of our webinar today but new rights of portability - there is an obligation in certain circumstances to conduct a Data Protection Impact Assessment within the company, there is an obligation to make notifications of personal data breaches to authorities and to data subjects. This is all feasible. Companies that take privacy into account properly will be ready for the GDPR but they must start well in advance because new processes must be implemented and put in place in the company. One important thing is that the fines can be scary for companies because it goes up to 20 million Euros or 4% of annual turnover and it is the highest amount that counts, so for big companies we are talking about 4% of worldwide turnover but it also gives a lot of leverage to privacy professionals to have more engagement by management because when you are trying to implement privacy safeguards and security safeguards and systems, the order of magnitude of fines is changing completely under the GDPR. So, this is a very interesting tool for privacy professionals - and this is how they must see it. The last point I would like to make with respect to these questions is the fact that the Article 29 Working Party and group of EU regulators are trying to help you with respect to the implementation of the GDPR. They have already issued some guidelines in December 2017 on several topics including the right to portability and data protection officers. They will produce guidelines on other topics - the hot topics - to help companies be ready for the GDPR in 2018. This will be done again and again until May 2018 when the GDPR is applicable.

Question No 2

US law focuses on whether identifiers are directly linked to data subjects, but EU law is focused on whether identifiers are **linkable** to data subject identities. The GDPR requires "appropriate technical and organizational measures to safeguard the rights and freedoms of data subjects" - does this mean that persistent identifiers are not permissible under the GDPR?

Gwendal Le Grand

The GDPR is a framework that explains under what conditions you can process personal data. It is not a ban on the processing of personal data. It says you can process personal data only under certain conditions. These conditions are the privacy principals described in the GDPR. One of these principals is security - where pseudonymization (defined as noted above under article 4(5) as requiring processing of data in a manner that can no longer be attributed to an individual data subject without the use of additional information provided such additional information is kept separately and subject to technical and organizational measures that ensure it cannot be attributed to an individual) can be implemented. The principal of the risked-based approach to data protection means you must understand what safeguards you must implement based in the types of risks that your processing is facing. It is not preventing the processing of personal data. The only thing that it is saying is that the GDPR does not apply to data that is anonymous.

Gary LaFever

The problem with persistent identifiers is who has access to those persistent identifiers and how likely are they to be subject to linkage attacks or re-identification via the mosaic effect. Like Gwendal says, the GDPR is not intended to stop the processing of personal data but rather you are supposed to put into place protective measures - both organizational and technical - to make unauthorized re-identification more difficult.

Mike Hintze

No type of data – persistent identifiers or otherwise – are barred from processing under the GDPR. Any type of data can be processed. But, if the data you process is personal data that is not de-identified in a significant way (like persistent identifiers) more severe restrictions under the GDPR will apply to you. Any level of effective de-identification shows that you are attempting to adopt the kinds of measures required under the GDPR. If you use an intermediate level of de-identification, like the Article 11 de-identification described earlier, you get additional relief. If you achieve the highest level of de-identification that meets the anonymization bar that Gwendal described, you get complete relief from the GDPR. It is a spectrum. There is nothing that is absolutely barred under the GDPR - it is a matter of the corresponding compliance obligations that apply to data based on the nature of the identifiability of that data.

Question No 3

Two questions were read together for a combined response

3(a) Can you suggest how to get budget for GDPR compliance in 2017 when senior management views GDPR as a 2018 issue?

3(b) As Chief Privacy Officer my title has a C in it but that does not mean I have a key to the C-Suite. The magnitude of liabilities and obligations under the GDPR are way out of synch with budget and authority I have in my position. How do I navigate the corporate labyrinth to make senior executives fully appreciate the magnitude of these issues?

Gwendal Le Grand

New processes that need to be implemented in companies take time. It takes time to be prepared adequately to make sure that companies are ready. Two important triggers that you can find in the regulation are Article 3 and Article 83. Article 3 is about the territorial scope and it says that the regulation applies to controllers and processors regardless of whether the processing is taking place in the EU or not. If you have an establishment or if you are targeting users in the EU, then the GDPR will be applicable to you. This means that the GDPR applies to many companies. If you are offering services in Europe, you need to take the GDPR into account. Article 83 is about the fines under the GDPR. Everyone has heard about this but if you go to your management and you say, "The risk that we run as a company if we are not prepared for the GDPR is a very large amount of money (up to the greater of 20 Million Euros or 4% of global gross revenues)" – this gives you a lot of leverage when you speak with management. It may not be a particularly nice way to speak with your management, but I guess it is very efficient.

Mike Hintze

This comes back to some of the things we talked about earlier with the first group of questions. You need to make the case that the time is now to be focusing on the GDPR. The types of things that need to put into place to show compliance with the GDPR are not things that you can just "flip a switch" or "turn on a dime." They require prioritization and action now and over the next year. So, laying that out and showing the types of things that need to be done – the type of architectural changes that may be required, the type of process changes, organizational changes and personnel training – these all take time. They require investments of time and money currently and if you are waiting until 2018 to do this, there just is not going to be enough time to get it done.

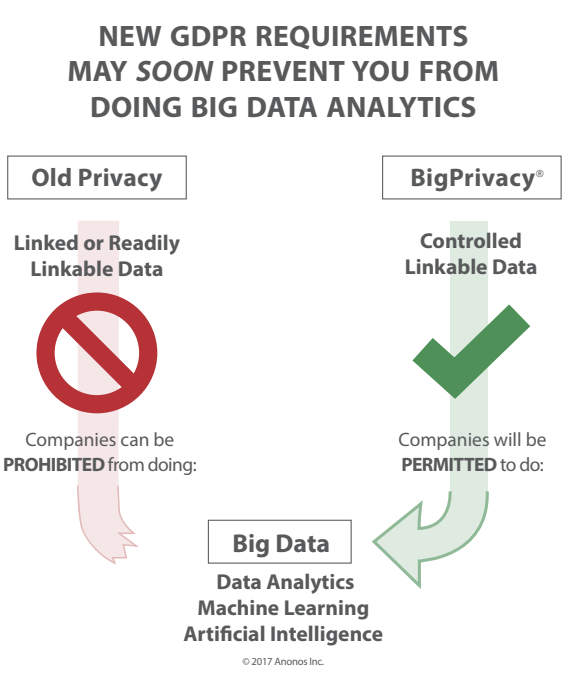
C. Summary

Gary LaFever

All companies are clearly at a tipping point. Companies can no longer do what they used to do 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 the data that they are making. I encourage you to look at the White Paper that Mike Hintze and I recently wrote on Big Data and Controlled Linkable Data (available at https://anonos.com/whitepaper). Readily Linkable Data and Linked Data as defined in the Big Data White Paper – the way they have been used in the past – can no longer be used the same way. You must have protective mechanisms in place and show that you are giving controls to data subjects and that you are respecting their rights. This requires new technical measures – data protection by default did not exist prior to the GDPR. As Gwendal just said, what changes on May 25, 2018 is as much about the fines as it is about new requirements. 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 to continue business processes so that everyone can be successful.

Note: Questions and answers not discussed live during the Webinar can be viewed at:
anonos.com/gdpr-webinar-faq

The following is a slide from the Webinar presentation



Controlled Linkable Data enables intelligent technical and policy solutions that deliver the benefits of data uses while avoiding the risks.

Jules Polonetsky, CEO
Future of Privacy Forum

Controlled Linkable Data tools minimize risk by de-linking and re-linking data to break the stalemate between responsible use and data obscurity.

Martin Abrams, Executive Director & Chief Strategist
Information Accountability Foundation

Excerpts from Hintze/LaFever White Paper available at https://anonos.com/whitepaper

IAPP Article, TRUSTe Blog, and TED Talk about Unlocking the Value of Data Analytics

IAPP Article - Maximizing data value while complying with GDPR may not be impossible

"What's been realized is this convenience of data processing came with a cost of the fundamental rights of data subjects. If viewed in the right light, the GDPR provides an answer to that," said Anonos CEO Gary LaFever during a recent IAPP web conference. "Data protection by default enables us to respect, honor and protect the fundamental rights of data subjects, while actually opening up new business opportunities."

Read the Article

TRUSTe Blog - Maximizing Data Utility Under GDPR

"Since so many businesses rely on big data analytics, as increasingly artificial intelligence, to fuel innovation and growth, it has become essential to know how to ensure compliance in a way that allows your data assets to be utilized."

"Hintze and LaFever present a compelling case for companies to proactively implement a robust technical approach to the GDPR's data protection by default requirements in order to both maximize data value and minimize compliance risk and liability."

Read the Blog

TED Talk

The principles of Anonos have been ahead of the curve on protecting privacy while maximizing the value of big data since 2012. Here are two quotes from a 6-minute TED Talk by Ted Myerson, Co-Founder of Anonos.

"Protecting privacy and maximizing the benefits from using Big Data do not have to be at odds"

"The benefit of infusing privacy and security into the data is that controlled sharing of highly personal and restricted data is made possible while protecting the privacy of individuals and minimizing risks."

Watch the 6 min TED Talk

Click to learn more about Anonos and how to unlock the value of Data Analytics under the GDPR

Or email us at:
LearnMore@anonos.com