New Requirements for Legal Analytics & Artificial Intelligence under the GDPR

May 2018

Anonos®

BigPrivacy® Unlocks Data

New Requirements for Legal Analytics & Artificial Intelligence Under the GDPR

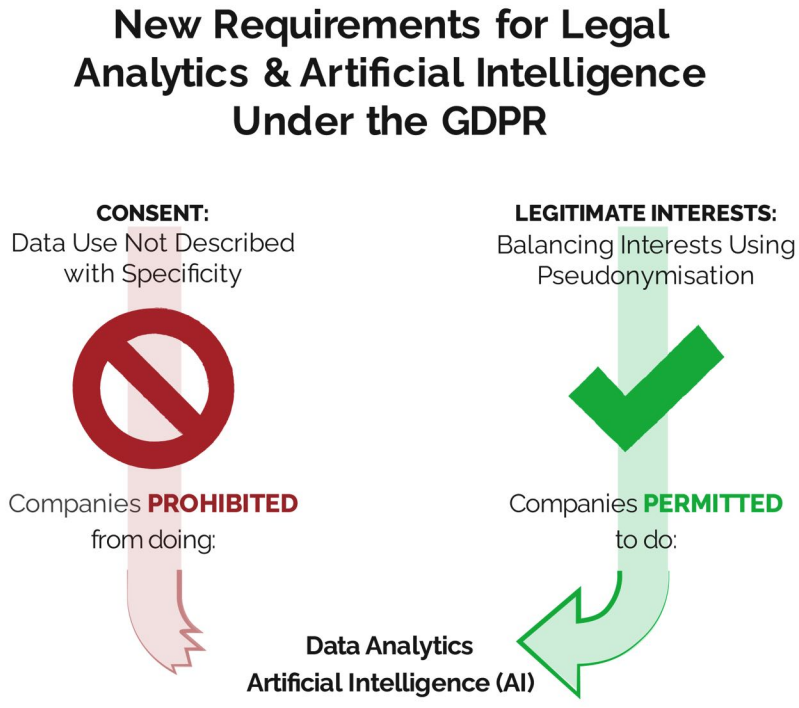
Summary

The way your organization has processed data for years – *even for decades* – may create new legal liability under the EU General Data Protection Regulation (GDPR). Organizations looking to maximize the value of data by leveraging advanced data analytics and artificial intelligence ("Analytics & Artificial Intelligence (AI)") should be aware of 3 points:

1. **Illegal Analytics & Artificial Intelligence (AI)** – without a new legal basis to ensure lawful rights to process personal data (which is no longer possible under the GDPR using consent if the processing cannot be described with specificity in advance), using that data for Analytics & Artificial Intelligence (AI) may produce unlawful results that expose organizations, their partners and customers to unexpected legal liability.

2. **Pseudonymisation** – while this word may be difficult to pronounce and spell, organizations must embrace and implement Pseudonymisation as defined under the GDPR across their enterprise to legally process Analytics & Artificial Intelligence (AI).[1]

3. **Dynamism** – new GDPR requirements for technically enforced "dynamism" overcome shortcomings of "static" data protection techniques that fail to adequately protect data subject rights when data is combined from multiple sources or used for various purposes. Examples under the GDPR include dynamic Pseudonymisation to defeat the Mosaic Effect and dynamic fine-grained, use-case specific controls to satisfy requirements for Data Protection by Design and by Default.



[1] The word "Pseudonymisation" is pronounced *Soo-don-uh-mah-zay-shuhn* and is generally spelled with an "s" (Pseudonymisation) in European countries and with a "z" (Pseudonymization) in the U.S. Under the GDPR, Pseudonymisation is defined as "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 organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person."

GDPR Impact on Analytics & AI Across the Globe

The GDPR is the new State of the Art for international data protection regulation with far-reaching impact on all organizations conducting business on a global scale.[2]

In today's data-driven economy, global organizations increasingly rely on Analytics & Artificial Intelligence (AI) to help achieve corporate growth, revenue and performance objectives. Historically, such organizations often relied on the consent (or agreement) of data subjects to long, legalistic take-it-or-leave-it contracts, "click-through" agreements and Terms of Service ("ToS") to secure legal rights to perform Analytics & Artificial Intelligence (AI) processing on personal data they collected. **Under the GDPR, data subjects can no longer legally consent (or agree) to data use that cannot be specifically and unambiguously explained at the time of consent. These GDPR requirements are impossible to satisfy for Analytics & Artificial Intelligence (AI) applications where successive analysis, correlations, and computations cannot be described with required specificity and unambiguity at the time of consent.**

The GDPR has no "grandfather provision" or "exemptions" allowing for continued use of historical data collected before the effective date of the GDPR using now legally non-compliant consent. Also, consent cannot be a condition for receiving a product or service – a data subject must have a genuine choice, or their consent is not freely given. If consent is not obtained in full compliance with GDPR requirements, DPAs have officially stated that "*consent will be an invalid basis for processing, rendering the processing activity unlawful.*"[3]

In addition to well-publicized fines and enforcement actions available to EU member state data protection authorities (DPAs), and the creation of new forms of liability under EU law,[4] global organizations should also be aware of the interests of other stakeholder groups in connection with their compliance with the GDPR. A high-level overview of some of the perspectives of these different stakeholder groups follows:

1. **External Auditors**; External auditors retained by organizations to conduct financial audits have an obligation to ensure that issued reports accurately represent the economic viability of the organizations. Given the adverse impact on business continuity of potential lost access to data and the magnitude of potential liability, if such auditing firms do not adequately describe the GDPR preparedness of a client in an audit, they may be liable for failing to sufficiently audit financial statements.

2. **Industry Regulators:** Industry regulators (e.g., banking, insurance, telecommunications, healthcare, etc.) are expected to require companies under their jurisdiction to confirm that they are not at risk of liability under the GDPR given the magnitude of potential financial exposure. For organizations depending on Analytics & Artificial Intelligence (AI) to help achieve corporate growth, revenue and performance objectives, use dynamic Pseudonymisation is required to confirm compliance.

3. **Boards of Directors:** There is a global trend toward directors being held personally liable for wrongdoings resulting from legislative changes and increased enforcement activity. An article notes, "The risk here is also exacerbated by the fact that the GDPR will impose much more stringent burdens on companies that process European Economic Area (EEA) citizens' data from May 2018. If significant penalties are imposed on companies under the GDPR – like the maximum penalties of 4% of an organization's worldwide turnover – shareholders, regulators and others may look to the board to ascertain what went wrong."[5]

---

[2] The GDPR is applicable to the processing of all "personal data" (information related to an individual data subject who can be directly identified using the data or indirectly identified by combining the data with other information) of any natural person located in the European Union (EU) or in the European Economic Area (EEA) at the time when processing occurs regardless of where a data controller or data processor is located if the processing relates (a) to the offering of goods or services (even if they are free) to a data subject located in the EU/EEA or (b) where the behavior of a data subject located in the EU/EEA is being monitored.

[3] See pg. 3 at http://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=51030

[4] Under the GDPR, member state data protection authorities (DPAs) have the power to: (a) impose fines of up to the greater of 20 million Euros or 4% of consolidated corporate global gross revenues for violating the GDPR and; (b) issue orders to stop or suspend data processing and data flows. Also, for the first time under EU law, the GDPR authorizes: (a) quasi-class-action lawsuits by representatives of individual data subjects (e.g., EU non-governmental organizations and advocacy groups); (b) recovery for non-monetary losses like damage to reputation, emotional distress, pain and suffering, and other injuries generally not recoverable under other jurisdictions' laws (e.g., in the US) for data protection, privacy or security claims; and (c) effective joint and several legal obligations among data controllers and processors up and down the data service "stack."

[5] See *Risks Facing Directors & Officers* at https://www.financierworldwide.com/roundtable-risks-facing-directors-officers-aug17/#.WwLxaVMvyuU

How Businesses Can Deal with Analytics & AI Under the GDPR

Before the GDPR, the primary burden of risk for inadequate data protection was born principally by data subjects, due to limited recourse against data controllers that collected and stored their data and lack of liability for data processors. However, this burden of risk is shifted by new obligations under the GDPR requiring data controllers and processors to protect the rights of individual data subjects. Given this change in circumstances, how do global organizations reconcile the growing importance of Analytics & Artificial Intelligence (AI) with the increasing complexity of lawful data use? The GDPR provides an answer in new requirements for technical and organizational safeguards that better protect personal data.
To lawfully process Analytics & Artificial Intelligence (AI) applications, and to legally use historical data collected using now legally non-compliant consent, new technical and organizational measures that help support alternate (non-consent) GDPR-compliant legal bases are required. Organizations must demonstrate that they have new technical and organizational safeguards to lawfully collect, use and store data for Analytics & Artificial Intelligence (AI) processing. Pseudonymisation and Data Protection by Design and by Default are newly defined technical and organizational safeguards under the GDPR that help enable fine-grained, risk-managed, dynamic use case-specific controls to reallocate risk for inadequate data protection from data subjects to corporate data controllers and processors.

GDPR Compliance Does Not Ensure Legal Analytics & Artificial Intelligence (AI)

Just because an organization is GDPR compliant does not mean that it can legally process Analytics & Artificial Intelligence (AI). Data protection technologies developed prior to the GDPR were not designed to support Analytics & Artificial Intelligence (AI) in a GDPR compliant manner. They are mainly intended to limit or prevent situations that expose an organization to potential penalty, liability, and third-party claims.

Traditional data protection technologies do not address the loss of consent as a viable legal basis for desired Analytics & Artificial Intelligence (AI) processing, they also create the perception that Analytics & Artificial Intelligence (AI) must be deleted and that lawful Analytics & Artificial Intelligence (AI) processing is no longer possible.

To maximize data value, new technical and organizational measures like GDPR compliant dynamic Pseudonymisation and Data Protection by Design and by Default are necessary to safeguard the privacy rights of data subjects in situations where consent is impractical, unavailable or unattainable.

Shortcomings of Traditional Security-Only Solutions

A conventional approach to improving data protection is to prescribe security upgrades. The problem with relying on this strategy by itself is two-fold. First, security solutions limit access to data by enforcing generalized access/no access controls to entire datasets, preventing people without permission from accessing any data, or granting access, to all of the data. Security-only solutions do not support fine-grained, risk-managed, use case-specific controls over what people can do with data once they receive access. Second, security technologies such as encryption, hashing, static or stateless tokenization, data masking, and related approaches, help to protect against unauthorized identification of data subjects using data that directly reveals the identity of a data subject, but do nothing to protect against unauthorized re-identification of data subjects by correlating data attributes to reveal identity via “linkage attacks” sometimes referred to as the “Mosaic Effect.”

Shortcomings of Traditional Privacy-Only Solutions

Before the GDPR, privacy was protected primarily using written contracts, “click-through” agreements and Terms of Service (“ToS”) that set forth what organizations would be authorized to do, or not do, with data. However, for nontechnical, non-preventive, policy-based measures to remain effective, controllers require resources and access to monitor compliance by all counterparties to contractual commitments. Such monitoring is typically unavailable or impractical to implement. Policy and contract-based measures also place the risk from inadequate data protection on data subjects, due to limited recourse against data controllers and data processors for privacy violations. Traditional, pre-GDPR technologies developed to safeguard privacy rights either work on a binary access/no access basis (e.g., data masking) or on an aggregated basis to support generalized statistics.

In today's changing regulatory landscape, these technologies fail to comply with new GDPR standards for modern digital processing or they cannot support business needs for increased access to personal data without the availability of consent. For example, combining and analyzing multiple data sets and incorporating unstructured data – processing which is at the core of the new digital economy – cause traditional privacy technologies to break down and prevent them from supporting GDPR compliant Analytics & Artificial Intelligence (AI) processing.

Traditional approaches to data protection employ outdated static approaches to pseudonymisation. As a result, supposedly pseudonymised data can be readily traced back to a data subject because the persistent or static pseudonymous tokens used for a given data element do not change. Searching for a random, pseudonymised string which repeats itself within or across databases can provide a malicious actor or interloper with enough information to unmask the identity of a data subject.

Consider the following simple example. A database contains a zip code value of "20500," so a static (or persistent) pseudonymous token of "6%3a8" is used to replace every occurrence where zip code 20500 is found. Due to advances in technology and threat-actor sophistication, such static pseudonyms can be readily linked back to individuals via the Mosaic Effect without requiring any access to additional information to reveal that pseudonym's value. An example of the Mosaic Effect is where three seemingly "anonymous" data sets that use persistent (static) pseudonyms – each composed of the zip code, age and gender of US citizens – were combined to identify up to 87% of the population of the United States by name.[6]

How to Legally Process Analytics and AI Under the GDPR

To achieve a sustainable competitive advantage and drive real value and revenue growth in today's increasingly regulated information economy, organizations must be on the forefront of maximizing the value of data by leveraging Analytics & Artificial Intelligence (AI); yet, at the same time, they must comply with new, more stringent requirements under the GDPR and similar evolving data protection regulations. Anonos® BigPrivacy® technology uniquely harmonizes two objectives which have previously been in opposition: it delivers data value through data use, sharing and combination while simultaneously enabling GDPR-compliant data protection for Analytics & Artificial Intelligence (AI).

Anonos BigPrivacy's patented use of dynamic (versus static) pseudonymisation is essential because the GDPR defines "Pseudonymisation" as requiring that personal data cannot be attributed to a specific individual without the use of additional information that is subject to protective technical and organizational measures. With BigPrivacy, a malicious actor or interloper can no longer determine that dynamically generated pseudonymous tokens belong or relate to the same data subject, let alone uncover a data subject's name or other identifying information. By replacing static unchanging pseudonymous tokens used in traditional approaches to data protection with patented dynamically pseudonymised tokens for each use and for each type of use, BigPrivacy dramatically mitigates risks of linkability and re-identification. BigPrivacy dynamic pseudonymisation enables fine-grained technical enforcement of policies and procedures under the purview of a data controller to safeguard the rights and interests of data subjects by technically enforcing GDPR principles of Pseudonymisation and Data Protection by Design and by Default.

BigPrivacy, the patented technology developed over the last six years by Anonos, is the State of the Art in Pseudonymisation & Data Protection by Design and by Default technology.

For more information on dynamic Pseudonymisation, contact us as LearnMore@anonos.com

[6] See http://dataprivacylab.org/projects/identifiability/paper1.pdf