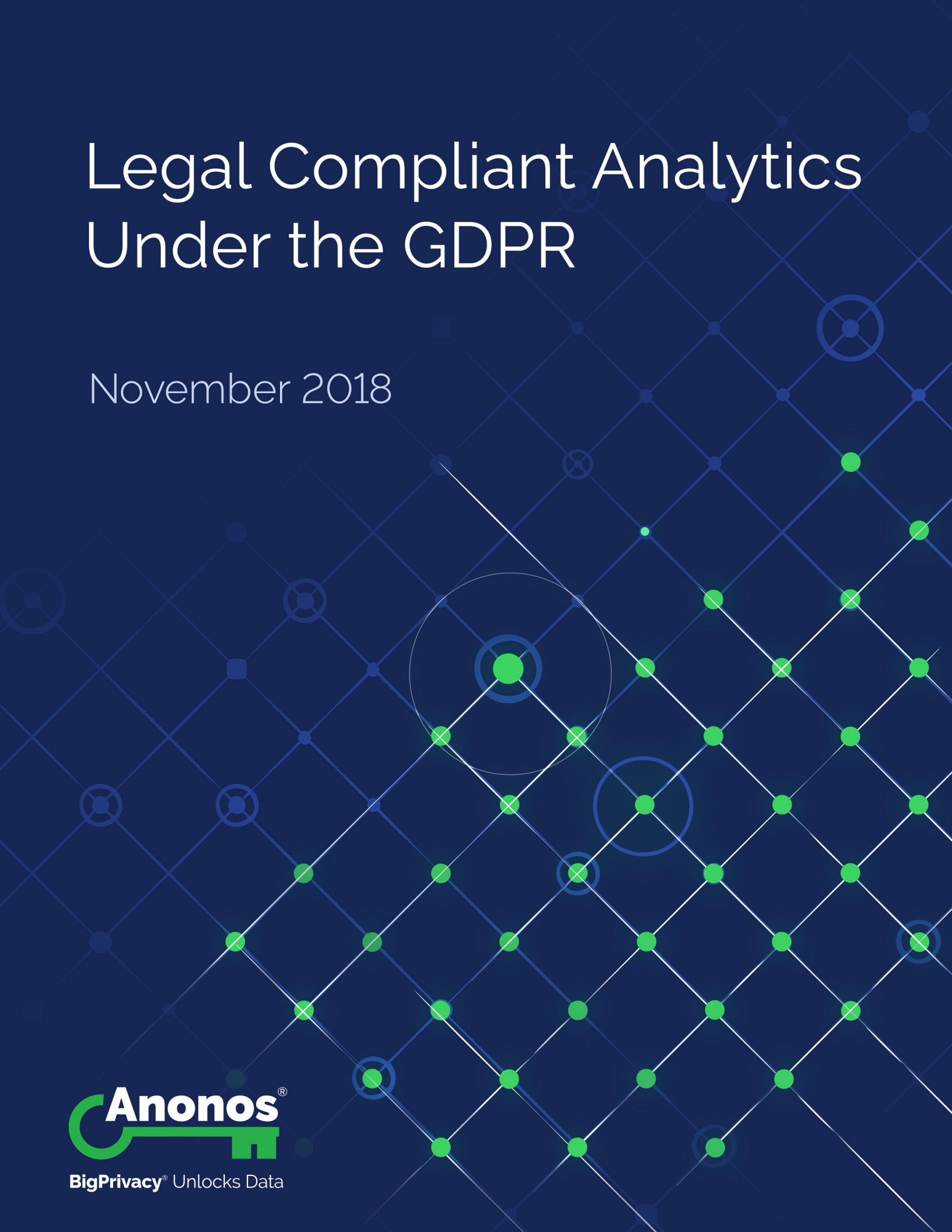
Legal Compliant Analytics Under the GDPR

November 2018



Anonos®

BigPrivacy® Unlocks Data

Anonos
BigPrivacy® Unlocks Data

Legal Compliant Analytics Under the GDPR

Summary

The way your organisation has processed data for years – ***even for decades*** – may now create legal liability under the EU General Data Protection Regulation (GDPR).

Organisations looking to maximize the value of data for advanced analytics, artificial intelligence, sharing, combination and repurposing ("Big Data Analytics & AI") should be aware of the following:

1. **Organisations Deleting Legacy Consent Data** – Organisations are deleting valuable data critical for Big Data Analytics & AI due to concerns over potential fines and injunctions ordering the immediate termination of illegal processing of data collected using non-compliant general consent ("Legacy Consent Data"). The GDPR has no "grandfather provision" or "exemptions" allowing for continued use of Legacy Consent Data. Also, consent to use data for Big Data Analytics & AI cannot be a condition for receiving a product or service – a data subject must have a genuine choice, or consent is not freely given and is unlawful.[1]

2. **Illegal Big Data Analytics & 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 Big Data Analytics & AI may produce unlawful results that exposes organisations, their partners and customers to legal liability.

3. **Pseudonymisation** – Organisations should embrace GDPR compliant Pseudonymisation[2] across their enterprise to process legal Big Data Analytics & AI. GDPR compliant Pseudonymisation technically enforces "dynamism" to support Legitimate Interest processing[3] by overcoming shortcomings of "static" data protection techniques that fail to adequately protect data subjects against unauthorized re-identification when data is combined from multiple sources or used for various purposes – known as the "Mosaic Effect."

4. **Certified Pseudonymisation Technology** – Anonos' patented SaveYourData® software is the only technology that has been certified as complying with legal and technical requirements for Pseudonymisation under the GDPR in accordance with the 'EuroPrivacy' certification scheme developed through a European research project co-funded by the European Commission and Switzerland.[4]

GDPR Impact on Big Data Analytics & AI Across the Globe

The GDPR is the new standard for international data protection regulation with far-reaching impact on all organisations conducting business on a global scale.[5]

In today's data-driven economy, global organisations increasingly rely on Big Data Analytics & AI to help achieve corporate growth, revenue and performance objectives. Historically, such organisations often relied on the general 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 Big Data Analytics & 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 Big Data Analytics & AI applications where successive analysis, correlations, and computations cannot be described with required specificity and unambiguity at the time of consent.

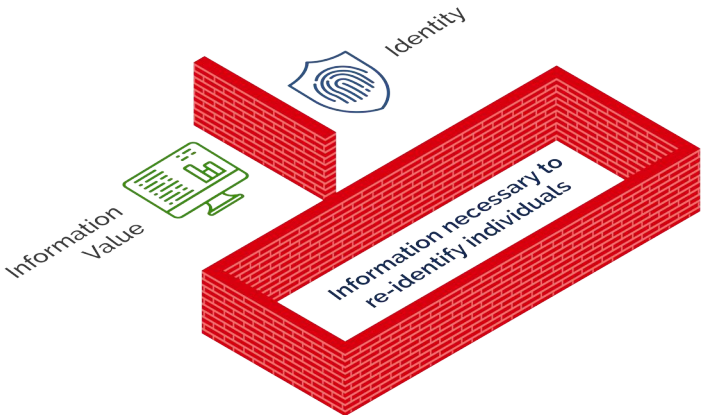
How Businesses Can Deal with Big Data Analytics & AI Under the GDPR

Before the GDPR, the primary burden of risk for inadequate data protection was born principally by individual 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 the reallocation of risk under the GDPR, how do global organisations reconcile the growing importance of Big Data Analytics & AI with the increasing complexity of lawful data use?

The GDPR provides an answer in new requirements for technical and organisational safeguards that better protect personal data. To lawfully process Big Data Analytics & AI, and to legally transform Legacy Consent Data, new technical and organisational measures that help support an alternate (non-consent) GDPR-compliant legal basis like Legitimate Interest are required. Organisations must demonstrate that they have new technical and organisational safeguards to lawfully collect, use and store personal data for Big Data Analytics & AI processing. GDPR compliant Pseudonymisation enables fine-grained, risk-managed, dynamic use case-specific controls to reallocate risk for inadequate data protection from individual data subjects to corporate data controllers and processors.

GDPR Article 4(5) Requirements for Pseudonymisation



1. Technical and organizational measures must separate the information value of data from the means of identifying data subjects.
2. Additional information must be required to re-link the data back to individual data subjects.
3. The additional information must be separately stored and made available only to persons with authorized access.

GDPR Compliance Does Not Ensure Legal Big Data Analytics & AI

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

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

To maximize data value, new technical and organisational measures like GDPR compliant dynamic Pseudonymisation 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 protecting data 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 generalised 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 among data sets to reveal identity via "linkage attacks" or the Mosaic Effect.**

Shortcomings of Traditional Privacy-Only Solutions

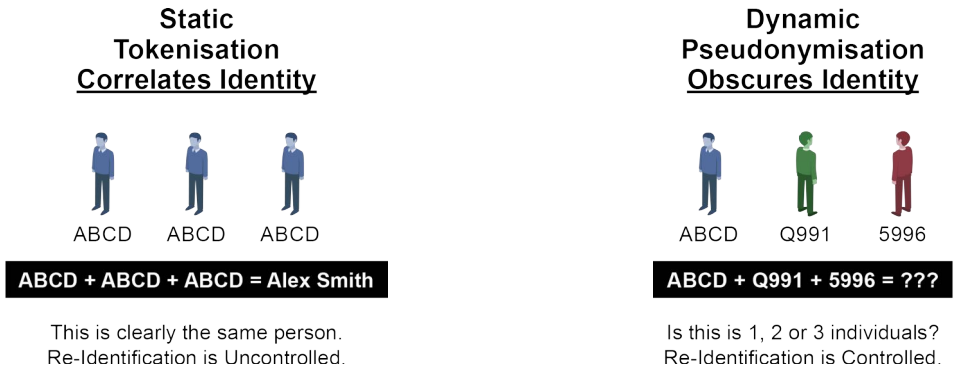
Before the GDPR, privacy was protected primarily using written contracts, "click-through" agreements and Terms of Service ("ToS") that set forth what organisations would be authorized to do, or not do, with data. However, for non-technical, non-preventive, policy-based measures to remain effective, controllers require resources and access to monitor compliance by all counterparties to contractual commitments. Such monitoring is typically unavailable or impractical to implement.

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 static techniques fail to comply with new GDPR dynamic standards and do not support business needs for increased access to personal data without having to rely on consent. **For example, dynamic data uses like 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 Big Data Analytics & AI processing.**

Traditional data privacy solutions employ outdated static approaches to protection. As a result, supposedly protected data can be readily traced back to a data subject because the persistent or static tokens used for a given data element do not change. Searching for a random, tokenised 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 postal code of "20500," so a static (or persistent) token of "6%3a8" is used to replace every occurrence where postal code 20500 is found. Due to advances in technology and threat-actor sophistication, such static tokens can be readily linked back to individuals via the Mosaic Effect to reveal that token's value. An example of the Mosaic Effect is where three seemingly "anonymous" data sets that use persistent (static) tokens – each composed of the postal code, age and gender of individuals – were combined to identify up to 87% of the population of the United States by name.[6]

Static vs Dynamic and the Risk of Re-Identification



How to Legally Process Big Data 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, organisations must be on the forefront of maximizing the value of data by leveraging Big Data Analytics & AI; yet, at the same time, they must comply with new, more stringent requirements under the GDPR and similar evolving data protection regulations. Anonos 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 Big Data Analytics & AI.**

Anonos' patented use of dynamic (versus static) data protection 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 organisational measures. With Anonos technology, 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 tokens used in traditional approaches to data protection with patented dynamically Pseudonymised tokens for each use and for each type of use, Anonos technology dramatically mitigates risks of linkability and re-identification.

Anonos technology enables fine-grained dynamic enforcement of policies and procedures under the purview of a data controller to safeguard the rights and interests of data subjects by technically enforcing the GDPR principle of Pseudonymisation.

CONCLUSION

Anonos BigPrivacy technology provides new legal and technology-based solutions to help enterprises address data privacy objectives brought about by recent GDPR and other regulatory data compliance requirements.

These new data privacy compliance and data minimisation requirements have had the impact of stalling Big Data analytics and AI projects at a time when demand for data intelligence is on the increase. These analytics and AI projects involve combining or sharing protected data sets across multiple divisions for cross-sell and up-sell profiling or sharing protected data sets with third parties for data validation and data enrichment processes.

The challenge is how to use this valuable data that is typically collected for primary purpose applications under a general consent clause or via contract for secondary use applications like Big Data analytics and AI.

- The GDPR limits the use of personal data collected to the initial purpose for which it was collected and prohibits use in secondary applications like Big Data analytics and AI, unless new technically enforced safeguards are put into place.
- Obtaining re-consent for each secondary use application is proving difficult and ineffective.
- Unless new technically enforced safeguards are put into place, data minimisation projects require personal data to be deleted so the data is unavailable for secondary use applications.
- Big Data analytics and AI projects are being restricted due to limitations on the transfer of personal data from different geographies to a centralised data lake.

Enterprises and other data protection solution vendors have had limited success complying with these complex data privacy requirements using encryption, static tokenization and differential privacy for complex Big Data analytics and AI use cases. These first-generation data protection technologies leave data sets vulnerable to unauthorized re-identification via the 'mosaic effect' and 'inference attacks' where the identity of a data subject can be revealed among data sets thereby compromising legal and regulatory compliance requirements.

Anonos BigPrivacy solutions uniquely help enterprises to satisfy new and evolving data privacy requirements by dynamically transforming personal data collected for primary purpose applications to support lawful secondary use applications under a new legal basis of Legitimate Interest. If implemented correctly, BigPrivacy enables an enterprise to repurpose data on its data driven journey to extract analytics and business intelligence value in a privacy respectful and compliant manner to maximize data value and improve competitive advantage.

For more information on certified GDPR compliant Pseudonymisation, contact us as LearnMore@Anonos.com

[1] See Article 29 Working Party Guidelines on Consent at https://iapp.org/media/pdf/resource_center/20180416_Article29WPGuidelinesonConsent_publishpdf.pdf

[2] 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 organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person."

[3] Legitimate Interest processing is defined under GDPR Article 6(1)(f) as processing "necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are 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." Legitimate Interest processing requires satisfaction of the following three-part test to show that data subjects' rights and interests are considered and protected: (1) Purpose Test: Are you pursuing a legitimate interest? (2) Necessity Test: Is the processing necessary for that purpose? (3) Balancing of Interests Test: Are technical and organisational safeguards, like GDPR compliant Pseudonymisation, in place so that the data subject's interests do not override the legitimate interest of the data controller or third party?

[4] See https://europrivacy.org

[5] 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 the citizenship of the person or where the 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.

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