ANONYMISATION UNDER THE GDPR

Anonos
BigPrivacy® Unlocks Data

BACKGROUND

What Constitutes Anonymous Data Under the GDPR?

The EU General Data Protection Regulation (GDPR) establishes a very high bar for what constitutes anonymous data, thereby exempting the data from the requirements of the GDPR:

> "…information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable."[1]

The European Data Protection Supervisor (EDPS) and the Spanish Agencia Española de Protección de Datos (AEPD) issued joint guidance related to requirements for anonymity and exemption from GDPR requirements. According to the EDPS and AEPD no one, including the data controller, should be able to re-identify data subjects in a properly anonymised dataset.[2]

This is in contrast to pseudonymised data which remains subject to the requirements of the GDPR as personal data. The Article 29 Data Protection Working Party (now the European Data Protection Board or EDPB) in its Opinion 05/2014 on Anonymisation Techniques ("WP29 Opinion 05/2014") provides reasons why pseudonymised data is not considered anonymised so as to be exempt from regulation:

- *"It is still possible to single out individuals' records if the individual is still identified by a unique attribute which is the result of the pseudonymisation function."*[3]
- *"Linkability will still be trivial between records using the same pseudonymised attribute to refer to the same individual. Even if different pseudonymised attributes are used for the same data subject, linkability may still be possible by means of other attributes."*[4]
- *"Inference attacks on the real identity of a data subject are possible within the data set or across different databases that use the same pseudonymised attribute for an individual."*[5]
- *"Simply altering the ID does not prevent someone from identifying a data subject if quasi-identifiers remain in the data set, or if the values of other attributes are still capable of identifying an individual."*[6] (i.e., indirect identifiers)
- *"If pseudonymisation is based on the substitution of an identity by another unique code, the presumption that this constitutes a robust de-identification is naïf and does not take into account the complexity of identification methodologies and the multifarious contexts where they might be applied."*[7] (i.e., use of static unchanging tokens, not protecting indirect identifiers)

ANONYMOUS DATA

How is This Reinforced?

Enforcement actions by EU Data Protection Authorities in 2019 further reinforced these points, for example:

1. In April, the Italian Data Protection Authority (Garante) ruled against Rousseau Association (as a Data Processor) with a finding that merely removing a telephone number when another persistent unique identifier still exists enabling indirect linking to data subject identities was inadequate protection of personal data.[8]
2. In March, the Danish Data Protection Authority (Datatilsynet) ruled against Taxa 4x35 (as a Data Controller) that merely deleting names while retaining persistent telephone numbers enabling linking to identities of data subjects was inadequate protection of personal data.[9]

SUMMARY

"Anonymisation" in Today's Data World

Recent well-publicized research by data scientists[10] at Imperial College in London and Université Catholique de Louvain in Belgium, as well as a ruling by Judge Michal Agmon-Gonen of the Tel Aviv District Court,[11] highlight the shortcomings of "Anonymisation" in today's big data world. Anonymisation reflects an outdated approach to data protection[12] that was developed when the processing of data was limited to isolated (siloed) applications prior to the popularity of "big data" processing involving the widespread sharing and combining of data. This is why the Israeli judge in the above-cited case highlighted the relevance of state of the art data protection principles embodied in the GDPR in her ruling that:

> *Increasing the technological capabilities that enable storing large amounts of data, known as "big data", and trading this information, enables the cross-referencing of information from different databases, and thus also trivial information such as location,* ***may be cross-referenced with other data and reveal many details about a person, which infringe upon his privacy.***
>
> *Given the scope of data collection and use of information, the matters of anonymisation and re-identification have recently become important and relevant to almost every entity in Israel – both private and public – which holds a substantial amount of information.*
>
> *Information technologies bring new challenges and ongoing privacy vulnerabilities. One of the solutions that has been discussed in recent years is that of privacy engineering (Privacy by design), i.e., the design of technological systems in advance, to include protection of privacy.*
>
> *A binding rule regarding privacy engineering was established in the European Union. Regulation for the Protection of Personal Data Article 25 of the GDPR General Data Protection Regulation (which came into effect in 2018) imposes* ***a duty on the data controller to implement appropriate and effective technological and organizational measures both at the stage of system planning and in the stage of information processing, in other words, requiring a process of privacy engineering.***

For data to be truly "Anonymous," the data must not be capable of being cross-referenced with other data to reveal identity. This very high standard is required because if data does satisfy the requirements for "Anonymity," it is treated as being outside the scope of legal protection provided under the GDPR. Why? Because of the very "safe" and protected nature of the data that actually satisfies the requirement of not being cross-referenceable or re-identifiable. The Israeli court in Disabled Veterans Association v. Ministry of Defense[13] highlighted that this is generally not the case in today's world of big data processing.

In today's world of big data processing, data that is held by a data controller may be readily linkable with data that is beyond the control of the controller thereby facilitating unauthorized re-identification and exposing:

(i) the data controller to potential liability;

(ii) data sharing partners of the controller to potential liability; and

(iii) data subjects to potential violations of their fundamental rights.

References

[1] GDPR Recital (26).

[2] See https://edps.europa.eu/sites/edp/files/publication/19-10-30_aepd-edps_paper_hash_final_en.pdf.

[3] See Article 29 Data Protection Working Party Opinion 05/2014 on Anonymisation Techniques, p.21.

[4] Id.

[5] Id.

[6] Id.

[7] Ibid, p.31.

[8] https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9101974

[9] https://www.datatilsynet.dk/tilsyn-og-afgoerelser/afgoerelser/2019/mar/tilsyn-med-taxa-4x35s-behandling-af-personoplysninger/

[10] https://www.nytimes.com/2019/07/23/health/data-privacy-protection.html?smid=nytcore-ios-share

[11] https://www.nevo.co.il/psika_html/minhali/MM-17-06-28857-22.htm

[12] https://www.timesofisrael.com/data-is-up-for-grabs-under-outdated-israeli-privacy-law-think-tank-says/

[13] Supra, note 11.