APPLYING THE EU GENERAL DATA PROTECTION REGULATION (GDPR) TO THE USE OF PERSONAL DATA TO TRAIN AN AI SYSTEM:

A follow-up to my note on a deeply defective and flawed decision of the Belgian data protection authority

by

Douwe Korff

Emeritus Professor of International Law, London Metropolitan University
Associate, Oxford Martin School, University of Oxford

16 April 2024

Douwe Korff
Emeritus Professor of International Law, London Metropolitan University
Associate, Oxford Martin School, University of Oxford

APPLYING THE EU GENERAL DATA PROTECTION REGULATION (GDPR) TO THE USE OF PERSONAL DATA TO TRAIN AN AI SYSTEM:

A follow-up to my note on a deeply defective and flawed Belgian decision

A few days ago, I published a highly critical note on a decision of the Dispute Resolution Chamber of the Belgian Data Protection Authority on the use, by a bank, of the transaction data of its customers to build direct marketing data models.[1]

The Belgian DPA's approach was based on a report by that same authority.[2] I have now learned that in fact this approach has been adopted more broadly. In particular, the French data protection authority, the CNIL, issued similar guidance a few years later.[3]

Although not based on a specific case (as the Belgian decision was), the CNIL too suggests that the "exclusive purpose" of the processing of personal data in the "learning phase" of the development of an AI system is "*to develop or improve the performances of [the relevant] AI system*". In other words, the CNIL, too, allows for the "*splitting off of 'phase one' [the learning phase] from the overall process*", and for then allowing the processing in this first phase on a separate legal basis from the overall purpose of the AI system.

But I maintain that, in relation to purpose-specific AI systems, **that makes no sense**. If an AI system is being developed for, say, the purpose of direct marketing, or to improve medical treatment of patients with a particular illness (or to improve the diagnosis of a particular illness), then surely the training of the system also serves that purpose: the training of the system cannot be separated from the ultimate purpose of the system; training such an AI system is not a purpose in itself. This has a major bearing on the issue of compatability (discussed in some detail in my previous note).[4]

Specifically, since "proximity marketing" (marketing one's own products and services to one's own customers) is compatible ("not incompatible") with the primary purpose for which personal data are processed (typically: to provide one's customers with the goods and services they asked for), then analysing one's own customer data to that end is also compatible with that primary purpose, and the same would apply to the AI-based creation of profiles of one's customers to underpin this marketing, *provided that*:

---

[1] Geschillenkamer van de Gegevensbeschermingsautoriteit, Beslissing ten gronde 46/2024 van 15 maart 2024, Dossiernummer: DOS-2019-05837 (hereafter: "**the decision**"), available at: https://www.gegevensbeschermingsautoriteit.be/publications/beslissing-ten-gronde-nr.-46-2024.pdf
My note on the decision, Applying the GDPR to AI-based marketing using banking data, 11 April 2024, is available at: https://www.ianbrown.tech/2024/04/11/applying-the-gdpr-to-ai-based-marketing-using-banking-data/

[2] Commissie voor de bescherming van de persoonlijke levenssfeer (CBPL), Big Data Rapport, 2018 (?), p. 34, available at: https://www.gegevensbeschermingsautoriteit.be/publications/big-data-rapport.pdf
See also the illustration and further detail on p. 35.

[3] CNIL, AI : ensuring GDPR compliance, 21 September 2022, original emphases, available at: https://www.cnil.fr/en/ai-ensuring-gdpr-compliance

[4] See my earlier note (footnote 1, above), sections 3.1 and 4.7.

- If only non-sensitive personal data are processed in the creation and training of the data model(s): the customers were informed of this in accordance with the GDPR[5] and did not object: that processing could then be based on the "legitimate interest" legal basis;

  But:

- If sensitive personal data are processed in the creation and training of the data model(s): the customers were informed of this in accordance with the GDPR and gave their explicit consent to the profiling: that processing cannot be based on the "legitimate interest" legal basis.

Moreover, since the making available (selling) of customer data by one company to another company for the latter's marketing ("third party marketing") is not compatible with the primary purpose for which personal data are processed (again typically: to provide the customers of the first company with the goods and services they asked for of that company), then AI-based analysing and the AI-based creation of profiles of those customers to underpin this third-party marketing is also incompatible with that primary purpose, and also requires their explicit consent (and that consent must be obtained separately from the consent for the processing for the primary purpose: see Article 7(2) GDPR).

The sleight of hand I accused the Belgian DPA of related to the fact that they would allow the use of any personal data for any of the above (proximity marketing based on non-sensitive data; proximity marketing based on sensitive data (!) and third party marketing based on either (!)) on the basis of the "legitimate interest" legal basis. As I said, if they reached that conclusion because they wanted to "help" the bank in question in its monetisation of its customers' transaction data, that was a deceit.

Somewhat, but not completely, different considerations apply in a health care context because of the special legal bases for such processing (noted below), but still also with distinctions between the following purposes, i.e.:*

(i) processing of patient data by a health institution (typically, a hospital) for the diagnosis and treatment of the patients (the primary purpose);

(ii) the processing of those same data by that same institution, by means of AI, for the secondary purpose of analysing the effectiveness of their diagnoses and treatment and improving them; and

(iii) the making available of those patient data to a third party:

  (a) for AI-based scientific research; and

  (b) for AI-based commercial purposes (e.g., to create and sell a fitness app).

* Of course, as the Belgian report and the CNIL guidance also note, it is often not easy to separate these purposes, but that should not mean that the distinctions can be ignored.

Processing of health care data (patient data) – which are inherently sensitive – is somewhat differently regulated than processing of personal data generally. They cannot be processed on

[5] See my earlier note (footnote 1, above), sections 3.2 and 4.5.

the "legitimate interest" legal basis (because they are sensitive). But they can be processed in particular on the following legal bases:

- the data subjects' (i.e., the patients') explicit consent (Article 9(2)(a));
- to protect the vital interests of the data subject or of another natural person where the data subject is physically or legally incapable of giving consent (Article 9(2)(c));
- on the basis of EU or EU Member States law if:
  - "*[the] processing is **necessary** for the purposes of preventive or occupational medicine, for the assessment of the working capacity of the employee, medical diagnosis, the provision of health or social care or treatment or the management of health or social care systems and services*" – provided that "*the data are processed by or under the responsibility of a professional subject to the obligation of **professional secrecy** under [EU or EU Member State law or medical sector rules]*" (Article 9(2)(h), read together with Article 9(3));
  - "*[the] processing is **necessary** for reasons of public interest in the area of public health, such as protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care and of medicinal products or medical devices*" – provided that "*[the law in question] provides for suitable and specific measures to safeguard the rights and freedoms of the data subject, in particular professional secrecy*" (Article 9(2)(i));

    And, if there is a derogation from *inter alia* the right to object:
  - "*[the processing is for] scientific or historical research purposes*" – provided that the law in question makes the processing "*subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject*" (Article 89 that further clarifies the data minimisations safeguards).

  (Such laws must of course also still always "*respect the essence of the right to data protection*", even though this is not explicitly mentioned in the above articles. Cf. Article 9(2)(g))
- on the basis of a contract between the data subjects (the patients) and a health professional – provided again that "*the data are processed by or under the responsibility of a professional subject to the obligation of **professional secrecy** under [EU or EU Member State law or medical sector rules]*" (Article 9(2)(h), read together with Article 9(3)); but in this case, no derogations can be made from data subject rights such as the right to object.

Note that in all the above cases the data subjects must still be **informed** of all the purpose**s** for which their data are to be processed, including any secondary and compatible purposes, and must be granted the right to **object** unless that is expressly set aside by law.

Notably, additionally, under the French law implementing and further specifying the rules in the GDPR, an authorisation from the CNIL is required for the establishment of AI data systems including systems that process health data – and such an authorisation will only be granted if

"suitable and specific"/"appropriate" safeguards are provided. This also, especially, applies to processing of health data for AI-based analysis and profiling.

The CNIL guidance refers to several cases that are relevant to the present issue:

> In the field of health, the CNIL has had the opportunity to give its opinion on the creation of health data warehouses. In recently published guidelines, it specifies the framework within which data can be collected and retained in a single database for a long period of time, as part of public interest missions and for subsequent research. (p. 3)
>
> As part of clinical research aimed at identifying explanatory variables for prostate cancer, the CNIL refused to allow a pharmaceutical laboratory to process data for the entire active patient population from the medical records of the various centres participating in the study.
>
> This active patient population in fact **contained hundreds of millions of records from individuals not suffering from prostate cancer** (and even records for women!). The desire to process this data, which is scientifically explained by the need for "true negatives" in order to effectively train a classifier, did indeed appear to be disproportionate to the purpose of the processing, and not necessary for the development of an effective AI system. (pp. 4 – 5, original emphasis)
>
> In the field of health, a clear distinction is made between the research phases, which require formalities to be completed with the CNIL (authorisation, compliance with a reference methodology, etc.), and the phases of use in a care pathway, which do not require any formalities to be completed with the CNIL. (p. 5)

The impression I get is that the CNIL feels it can address the issue of the proportionality and permissibility – and thus, in its view, of the legal basis – of AI-based profiling separately from the legal basis for the broader processing, precisely because it can impose its own conditions (safeguards) or refuse to allow the secondary processing if it is not content with the proposed AI-based activities. The EU legislator appears to have taken a similar view with regard to authorisation of secondary processing by law: the EU or any Member State can authorise AI-based profiling provided it makes it subject to the imposition of serious conditions and constraints.

But *outside of such special DPA-issued or law-based authorisations*, I maintain my point: **the legal basis for the creation and training of AI models to be used for some specified purpose (be that direct marketing or health care)* must be the same legal basis as the one that covers the purpose for which those models are to be used – because the processing to create and train the models serves that ultimate purpose.**

* I will come to General Purpose AI below.

In other words, *outside of such special DPA or law-based (conditional) authorisations*:

- if a health care provider wants to use the health data of its own patients for the secondary purpose of AI-based analysing of the effectiveness of its own diagnoses and treatment and improving them (purpose (ii), above), it must seek their <u>explicit consent</u> for this because the data are sensitive and the "legitimate interest" legal basis cannot be relied on: this may be a purpose that is compatible ("not incompatible") with the primary purpose

(treatment: purpose (i), above), but it still needs the same legal basis as the one that is (rightly) invoked for the primary purpose;

- explicit consent is also required if the health care provider wants to disclose (sell) the patient data to a third party (typically, an academic research institution) for non-commercial scientific research (purpose (iii)(a), above): the data are sensitive and the "legitimate interest" legal basis cannot be relied on; this too may be a purpose that is compatible ("not incompatible") with the primary purpose (treatment: purpose (i), above), but it still needs the same legal basis as the one that is (rightly) invoked for the primary purpose (and in this, the research institution must moreover abide by the conditions and limitations imposed by EU or national law or scientific sectoral rules that contain the "appropriate safeguards" mentioned in Article 89 GDPR); and
- the same also applies if the health care provider wants to disclose (sell) its patient data to a third party (typically, a pharmaceutical firm) for the creation and training of AI data models to be used for the third party's commercial purposes (purpose (iii)(b), above): this too requires the patients' explicit consent.

The misleading suggestion that secondary processing of personal data – and even sensitive data – for the training of AI-based data models (i.e., profiles) can be based on a different legal basis than the one that applies to the processing for the purpose for which the models/profiles are going to be used is wrong. The only exception or qualification to this is the situation just discussed: if a relevant EU or EU Member State law or an authorisation from a DPA under such a law specifically allows the data modelling/profiling – but in that case the law or authorisation will, and must, also set out crucial conditions and safeguards.

- **Some further comments *re* general purpose AI systems**

The above may not apply to General Purpose AI (GPAI) systems that are defined in the recently adopted EU AI Act as:

> an AI system which is based on a general-purpose AI model, that has the capability to serve a variety of purposes, both for direct use as well as for integration in other AI systems.
>
> (Article 3(66) –

with the latter (a GPAI model) being defined as:

> an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are released on the market.
>
> (Article 3(63))

In other words: GPAI systems and models are not developed for any specific, pre-determined purpose, but for use in support of all sorts of purposes.

But that of course makes them all the more risky. Indeed, they can pose "*systemic risk*" to "*public health, safety, public security, fundamental rights, or the society as a whole*" (Article 51(1)(a) read with Article 3(66) of the AI Act).

This raises two issues under the GDPR. First of all, can it really be said that the processing of personal data involved in the creation and training of GPAI models is done for a "*specified, explicit and legitimate purpose*" – as all processing of personal data must be (Article 5(1)(b)) – when in fact no purpose at all is specified when they are created? If the answer were to be in the negative, then no use of personal data to do this would be allowed under the GDPR.

But perhaps developments have gone too far to stop altogether. In that case, in my opinion, in terms of the GDPR, GPAI systems and models should be regarded as suitable to either regulation by law (with the law imposing "appropriate safeguards"), or under conditions spelled out in an authorisation from the relevant (competent) data protection supervisory authority issued under the law (with those conditions imposing the relevant safeguards – in the way the CNIL has done in the examples mentioned in its guidance, quoted above).

And even then, the deployment of any GPAI system or model for any specified purpose would, in my opinion, still inherently carry a "high risk", and should therefore not be done until a data protection impact assessment is carried out, resulting in appropriate safeguards for such a deployment (see Article 35 GDPR). In fact, the AI Act expressly makes clear that the transparency details that are to be provided on any high-risk AI system under Article 13 of that Act must also, "where applicable", be used to carry out a DPIA (Article 26(9) of the AI Act).

Conclusion

As a matter of principle and law, the creation and training of AI models/profiles for a specific purpose (be that direct marketing or health care) must be based on the legal basis relied on for that ultimate purpose.

The fact that the creation and training of the models/profiles is a "first phase" in a two-phase process (with the deployment of the models/profiles forming the "second phase") does not alter that.

However, as an exception to this, under the GDPR, the processing can also be authorised by law or by means of an authorisation issued by a DPA under the relevant law (as in France), provided the law or DPA authorisation lays down appropriate safeguards. That is the only qualification I accept to the above principle.

The creation and training of General-Purpose AI systems and models, which by definition are not developed for any pre-specified purpose but for use for a wide range of purposes, arguably breaches the purpose-specification principle set out in Article 5(1)(b) GDPR. They are in my opinion best suited for (strict) regulation by law or, as in France, by DPA authorisations, laying down appropriate safeguards.

And any deployment of a GPAI system for a specific purpose should still be subject to a data protection impact assessment (DPIA).

Douwe Korff (Prof.) Cambridge (UK), 16 April 2024