Health Data ('Data Concerning Health') Inferences – an Unexplored Volcano Eager to Erupt Post COVID-19 Crisis

Aligning Inferences with GDPR Principles?

Alina Škiljić

Winner of the 2020 ELI Young Lawyers Award

Executive Summary

We live in an unprecedented time – 2020 will undoubtedly be remembered for the COVID-19 crisis. While hardly any legal area has been unaffected, efforts to combat the virus have also initiated a discussion about certain EU data protection questions. DPAs agree that personal data protection must be ensured even in these exceptional times, especially when it comes to the generally prohibited processing of health data. The private sector is under a great deal of scrutiny when it comes to processing health data and it is unlikely that appropriate legal bases could be presented. Notwithstanding the challenge to control the processing activities performed on information clearly considered as personal data, inferences might become very influential in shaping business activities of the private sector during and post the COVID-19 crisis. This paper aims to provide an overview of the most pressing risks related to drawing inferences on health data, using the example of how the hotel industry might create inferences of guests' COVID-19 status based on the data of their origin of travel. The main issues that are presented deal with how (inaccurate) inferences might result in unjustified different treatment of data subjects and diminish the key GDPR principles, even when not created by sophisticated algorithms. The paper argues that providing the private sector with clear guidance in the form of a 'unique' necessity test and opening a comprehensive dialogue between regulators and the industry in light of the COVID-19 crisis could ensure that data processing is performed in accordance with applicable legislation. Looking beyond the COVID-19 crisis, precise guidelines on inferred data processing should be adopted to provide a framework for improving the overall data protection scheme.

List of Abbreviations

Article 29 WP	Article 29 Data Protection Working Party
CJEU	Court of Justice of the European Union
DPA	Data Protection Authority
EDPB	European Data Protection Board
EU	European Union
GDPR	General Data Protection Regulation
ICO	Information Commissioner's Office

Table of Contents

1. Introduction: Inferences as Personal Data – Long Standing Enemies on Shaky Ground

Considering data protection, inferences may be defined as

> *Information relating to an identified or identifiable natural person created through deduction or reasoning rather than mere observation or collection from the data subject*.[1]

The debate on inferences has well-developed roots, while their legal regulation is still not unified. Inferences present a high risk for privacy – they are not personal data voluntarily provided by data subjects, but rather data which has been created from the (voluntarily) provided data or used by data controllers or third parties.[2] Their dangerous aspect is low verifiability, even more so because they (might) result in important decisions.[3]

Although both the academics and regulators agree on the privacy-invasive nature of inferences, inferences still lack precision in their legal treatment.[4] To this extent, a common question is what constitutes personal data – and subsequently what rights do data subjects have.[5]

---

[1] Wachter, S. and Mittelstadt, B., A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI [2018], *Columbia Business Law Review* 2019(2) p 22 (Wachter, Mittelstadt *CBLR 2018*), <available at SSRN https://ssrn.com/abstract=3248829>. Also, see Hu, R. and Stalla-Bourdillon, S. and Yang, M. and Schiavo, V. and Sassone, V. Bridging Policy, Regulation, and Practice? A Techno-Legal Analysis of Three Types of Data in the GDPR, also published in Data Protection and Privacy: The Age of Intelligent Machines [2017], Leenes Rosamunde van Brakel, R. and Gutwirth, S. and De Hert, P. (eds) (Hart Publishing, 2017) <available at SSRN https://ssrn.com/abstract=3034261>.

[2] The differentiation between 'data provided by data subject' and 'inferred' or 'derived' data is presented in the Article 29 WP, Guidelines on the Right to Data Portability, 16/EN, p 9–11 [2016], <available at https://ec.europa.eu/newsroom/document.cfm?doc_id=44099>. According to these guidelines, inferred data and derived data are created by the data controller on the basis of the data provided by the data subject. As an example, Article 29 WP argues that the outcome of an assessment regarding the health of a user cannot in itself be considered as data provided by the data subject. The same conclusion can be found in Article 29 WP, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, 17/EN [2016]

[3] Wachter, Mittelstadt *CBLR 2018*, p 90.

[4] For the overview and analysis of the academic discussion on the classification of inferences as (sensitive) personal data see Wachter, Mittelstadt *CBLR 2018*, p 74-77. In addition, the question of classification of inferences appeared several times before the Court of Justice of the EU (CJEU). The jurisprudence of CJEU dealing with what constitutes a personal data when it comes to inferences is not conciliated, ie judgements differ in their interpretations of personal data. For example, when interpreting whether the opinions and assessments related to data subjects can be constituted as personal data in Joined Cases C–141 & 372/12, *YS, M and S v. Minister voor Immigratie, Integratie en Asiel*, [2014] E.C.R. I-2081 the CJEU interpreted personal data restrictedly. Conversely, in Case C-434/16, *Peter Nowak v. Data Prot. Comm'r*, [2017] E.C.R. I994, the CJEU took a stand that opinions and assessments fall in the scope personal data. For further discussion on the jurisprudence of the CJEU regarding the scope of personal data see Wachter, Mittelstadt *CBLR 2018*, p 29-50.

[5] For example, the Article 29 WP argued back in 2007 that data which is '*likely to have an impact on a certain person's rights and interests*' is sufficient for it to be treated as personal data, Article 29 WP opinion 4/2007 on the concept of personal data, 07/EN, p 11, [2007]. For further discussion of this matter, see Purtova, N., The law of everything. Broad concept of personal data and future of EU data protection law, [2018], *Law, Innovation and Technology*, 10:1, p 40-81, <available at https://doi.org/10.1080/17579961.2018.1452176>. See also Wachter, Mittelstadt *CBLR 2018*, p 34-35.

To illustrate the example of an inference 'inspired' by COVID-19, imagine a situation where someone is considered virus-positive or -negative or grouped under the probability of being positive (eg high or low risk) – based on their origin of travel. This would be a subjective assessment (a type of inference), as it involves inferring a non-observed characteristic (and a sensitive one – health data[6]) of the subject from data already held.[7] Not only could such processing be unlawful under the GDPR – but the created inferences might also be used for making decisions on data subjects. An inference does not even need to be correct – it would be personal data if completely inaccurate.[8] Accordingly, making decisions on inaccurate data seems even more intrusive.

Inferences are undoubtedly a hot topic in the age of Artificial Intelligence, Internet of Things and similar technology;[9] however, this paper aims to illustrate that personal data are under threat not only from inferences constructed by sophisticated algorithms, but likewise from ones which may be constructed in daily business practices and with a small amount of 'input' data – and ones which might be completely inaccurate and still significantly impact data subjects.

[6] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), published in *Official Journal of the European Union*, L 119, [4 May 2016] (GDPR). Note that the GDPR uses the terminology '*data concerning health*', however for the purposes of this paper the term '*health data*' will be used throughout.

[7] Wachter, Mittelstadt *CBLR 2018*), p 27-28.

[8] Article 29 WP 4/2007 on the concept of personal data, 07/EN, p 11, [2007]. Also Wachter, Mittelstadt *CBLR 2018*, p 76-77. As authors argue, accuracy and verifiability of inferred data do not diminish the impact the processing might have on data subjects.

[9] For analysis and discussion of 'technology-created inferences' see Wachter, S., Normative challenges of identification in the Internet of Things: Privacy, profiling, discrimination, and the GDPR, [2018], *Computer Law and Security Review* 2018(3), <available at: https://doi.org/10.1016/j.clsr.2018.02.002>. For an interesting view on how to solve the inference threat see Dzięgielewska, O., Anonymization, tokenization, encryption. How to recover unrecoverable data, [2017], *Computer Science and Mathematical Modelling*, 2017(6), p 9-13.

2. Health Data and the Private Sector: Can We Have Your Contact, Credit Card and Medical Record?

> *Personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status*.[10]

Health data can be about an individual's past, current or future health status; it includes any data which reveal anything about the state of someone's health.[11] This surpasses the term 'medical' and it is irrelevant whether it is 'ill health' data or 'healthy health' data.[12]

Health data can thus include a wide range of personal data; COVID-19-related health data would undoubtedly be whether a data subject has symptoms or has tested positive (or negative). It may also include information as to whether a data subject has been self-isolated or their body temperature.[13] Inferences or even (inaccurate) assumptions on someone's health can be considered health data – when conclusions are drawn about someone's health, regardless of their reliability, these conclusions are to be treated as health data.[14]

2.1. Applicable Legal Basis for Health Data Processing – and is There One for the Private Sector?

Health data is, by its nature, much more sensitive than eg e-mail; it exposes individuals to risks to their fundamental rights and freedoms. GDPR considers health data as a sensitive category

[10] GDPR, Art. 4(15). GDPR further refers to data concerning health in Recitals (35), (45), (52-54), (63), (71), (75) and (112).

[11] Further examples on what is health data can be found in the Information Commissioner's Office (ICO) guide on special category of data, <available at https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/special-category-data/what-is-special-category-data/>, accessed on 25 April 2020 (ICO guide on special category of data,). For example, health data can be any information on injury, disease, disability or disease risk, including medical history, medical opinions, diagnosis and clinical treatment; medical examination data, test results, data from medical devices, or data from fitness trackers; etc.

[12] Annex to the letter of Article 29 WP responding to a request of the European Commission to clarify the scope of the definition of health data in connection with lifestyle and wellbeing apps, [5 February 2015], <available at https://ec.europa.eu/justice/article29/documentation/otherdocument/files/2015/20150205_letter_art29wp_ec_health_data_after_plenary_annex_en.pdf>(Article 29 WP, Annex 2015).

[13] For example, see Belgian Data Protection Authority guidelines for employees' data processing during COVID-19, <available at https://www.autoriteprotectiondonnees.be/covid-19-et-traitement-de-donn%C3%A9es-%C3%A0-caract%C3%A8re-personnel-sur-le-lieu-de-travail>, accessed 20 April 2020.

[14] Article 29 WP, Annex 2015, p 4. In addition, Article 29 WP stresses in its Annex that when conclusions are drawn about a person's health status or health risk (irrespective of whether these are accurate or not and legitimate or not, or otherwise adequate or inadequate), such conclusions are to be treated as health data. Likewise, ICO in its guide on sensitive data argues that, in case of any form of profiling which infers, for example, ethnicity, beliefs, politics, health risks, sexual orientation or relationship status (which are all sensitive data, as well as health data), if such inferences are deliberately created then such processing should be considered as processing special category data irrespective of the level of statistical confidence. For detailed discussion of this matter see Malgieri, G. and Comandé, G., Sensitive-By-Distance: Quasi-Health Data in the Algorithmic Era, [2017], *Information, Communication and Technology Law*, 2017 (3), <available at SSRN: https://ssrn.com/abstract=3020628>.

of data, which benefits from much broader protection and the processing of which is, by a general rule, prohibited.[15] The processing of health data is only exceptionally permitted, if one of an exhaustive list of legal bases exists. In other words, controllers need, in addition to a 'basic' legal basis[16], one of the legal bases for processing sensitive personal data.[17]

It seems that, since the start of the COVID-19 outbreak, the public and the private sector massively considered that the processing of health data falls either under reasons of substantial public interest or under reasons of public interest in the area of public health.[18] The EDBP clarified that data protection rules do not compete with measures against COVID-19.[19] It is indisputable that data plays a crucial role in containing the spread of COVID-19; however, not every data processing can be so easily justified. DPAs rapidly followed with clarification of what legal basis might be appropriate – and what processing might be allowed, proportional and necessary. DPAs' guidelines do not currently deal with all possible health data processing situations; rather, they deal with processing of health data by public (health) authorities and by employers. When the processing of health data is not performed by health authorities, EU DPAs interpret the relevant provisions of the GDPR in a more restrictive manner. [20]

Having in mind the above, it seems straightforward that justifying the processing of health data by the private sector (eg accommodation) as an adjustment to COVID-19 would be

[15] Article 9 (1) GDPR.

[16] Article 6 (1) GDPR states that data processing is lawful only if and to the extent that at least one of the thereby listed legal bases applies.

[17] Article 9 (2) GDPR lists ten legal bases for processing of special (sensitive) category of personal data.

[18] Article 9.2 (g) GDPR reads 'processing is necessary for reasons of substantial public interest, on the basis of Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject'. Article 9 (i) GDPR reads: '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, on the basis of Union or Member State law which provides for suitable and specific measures to safeguard the rights and freedoms of the data subject, in particular professional secrecy'.

[19] Statement by the EDPB Chair on the processing of personal data in the context of the COVID-19 outbreak as of 16 March 2020, <available at https://edpb.europa.eu/news/news/2020/statement-edpb-chair-processing-personal-data-context-covid-19-outbreak_en>, accessed 25 March 2020.

[20] See Etteldorf, C., COVID-19 Special EU Member State Data Protection Authorities Deal with COVID-19: An Overview, [2020], *European Dana Protection Law Review*, 2020(2) preview <available at https://www.lexxion.eu/wp-content/uploads/2020/03/COVID-19-Special-Data-Protection-Authorities-Deal-with-COVID-19.pdf>, accessed on 20 April 2020. Access to all publicly available guidelines and statements of EU Data Protection Authorities (DPAs) is available at: < https://iapp.org/resources/article/dpa-guidance-on-covid-19/ >., accessed on 20 April 2020. Contrary to the majority of DPAs' stands expressed in guidelines and statements, the World Health Organization (WHO) in its Operational considerations for COVID-19 management in the accommodation sector states for the tourism accommodation sector that '*it is advisable to monitor potentially ill guests in the establishment*'. Operational considerations for COVID-19 management in the accommodation sector, Interim guidance, [31 March 2020], <available at https://apps.who.int/iris/bitstream/handle/10665/331638/WHO-2019-nCoV-Hotels-2020.1-eng.pdf>, accessed 2 April 2020.

insufficient. It would indeed be hard to imagine that DPAs would consider it proportional, necessary, and justified under the GDPR that, eg a hotel processes guests' medical records, performs temperature measurements or greets guests with 'health' questionnaires, although some of these might be proportionate under certain conditions – for instance, temperature measurements might be lawful if a hotel measures guests' temperature daily during their stay without storing this data (ie keeping records of guests' temperatures) or linking it to other personal data (ie identifying an individual guest, unless the temperature is high)[21], as it can be argued that, if so, temperature would not be a 'personal data'[22]. However, it cannot be, under any circumstance, expected (or should be allowed) that the private sector can ask '*Can we have your contact, credit card, and medical record?*' Again, it would be equally hard to imagine that the private sector would continue operating without any additional level of carefulness as if COVID-19 had never been around. Their reputation might be at stake – no organization wants to be seen to diminish data protection and hardly anyone would benefit from being labelled as a place where the virus spreads. If they mostly cannot lawfully process health data, does this open a space for the industry's autonomy to assess people relatively freely, regardless of laws that limit this freedom? This is where inferences might appear – and the ones that the private sector can draw are even more dangerous.[23]

3. 'Hidden' Health Data Processing in the Private Sector – Inferences Ready to 'Explode' Post COVID-19 Outbreak?

What if it is possible to infer someone's health status from the origin of his or her travel? One can surely make assumptions on that basis – and they do not even have to be accurate to be an inference. Results of COVID-19 tests (or answers to questions like '*Are you COVID-19 positive*? *Do you have any symptoms?*') are certainly sensitive (health) data. Location data (eg origin of travel), conversely, is not – and eg hotels already process this type of information for lawful purposes. This could easily be used for creating (unlawful) inferences about the guests' COVID-19 status (health data).

[21] See, for instance, opinion of the Croatian Data Protection Agency („AZOP') available at: <https://azop.hr/misljenja-agencije/detaljnije/sustav-za-mjerenje-tjelesne-temperature-putem-termalnih-kamera>, accessed 25 August 2020.
[22] Article 4 (1) GDPR defines personal data as 'any information relating to an identified or identifiable natural person ('data subject')…'.
[23] Wachter, Mittelstadt *CBLR 2018,* p 49.

The main challenge before regulators (and what should be one of their priorities) is how to stop the private sector from making assumptions on data subjects' health during and post the COVID-19 crisis and ultimately from unlawful health data processing.

3.1. Practical Observations of Possible (Unlawful) Health Data Inferences

It is indisputable to which lengths in today's global technology inferences could go – significant harm can be done by making assumptions based on a single piece of data. The ICO stated that it may be possible to infer details about someone which fall within the special categories of data; whether this counts as special category data and triggers Article 9 of the GDPR depends on how certain that inference is, and whether it is deliberately drawn.[24]

To illustrate this with a possible real-life scenario – a hotel observes reservations and includes information on incoming guests in its database (eg names, contact, origin of travel, pre-payment, etc.). A hotel considers the origin of travel as a parameter for the likelihood of COVID-19 infection (eg guests coming from highly affected areas are classified as possibly infected). It decides to group the guests based on their origin of travel into categories as 'high', 'low', 'medium' risk.[25] Although health data is not easily spotted in this probability assessment as, *de facto*, a hotel processes origin of travel, a hotel effectively makes assumptions, thus draws its own conclusions of a guest's health status – this conclusion (or even a suspicion (of illness)) may be seen as a (sensitive) inference.

Based on ICO's clarification, two things should be assessed when determining whether an inference constitutes sensitive data. If relevant information can be inferred with a reasonable degree of certainty, then it is likely to be sensitive data.[26] If we, for example, take into account

---

[24] ICO guide on special category of data, available at: <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/special-category-data/what-is-special-category-data/>. For additional observations of when inferences might constitute (sensitive) personal data, see <http://paulvines.com/2018/02/21/gdpr-and-latent-variables.html>, accessed on 1 April 2020. For the analysis of academic discussion on the classification of inferences as sensitive data see Wachter, Mittelstadt *CBLR 2018*, p 74-77. The authors stress that '*the classification of inferences as sensitive data potentially depends on two conditions: (1) the intention of inferring sensitive attributes, and (2) the reliability of the data in question for inferring sensitive attributes.*'

[25] The hotel (accommodation) industry is taken as an example as one of the industries highly affected by the COVID-19 crisis. See, for example Ross, D., The Economic Impact of COVID-19 on the Hotel Industry, Hospitality Technology, [2020], available at: <https://hospitalitytech.com/economic-impact-covid-19-hotel-industry>, accessed on 20 April 2020. Notwithstanding the recent developments, influenced by the COVID-19 crisis, the hotel industry has generally been an area of interest of DPAs. The overview of fines and penalties which DPAs within the EU have imposed under the GDPR is available at <https://www.enforcementtracker.com/>.

[26] ICO guide on special category of data, <available at https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/special-category-data/what-is-special-category-data/http://paulvines.com/2018/02/21/gdpr-and-latent-variables.html>.

the population of Spain – 46,751,717[27] – then the mere fact that a hotel guest is coming from Spain would not give a reasonable degree of certainty on the guest's COVID-19 status and would consequently not be health data but a possible inference. Moreover, if there is no reasonable degree of certainty, but rather just a possible inference or an 'educated guess'[28]- then it is not sensitive data – even if the inference or guess turns out to be right.[29]

At first sight, these observations might seem like a solution to circumvent the GDPR rules. However, ICO further explains that these 'exceptions' apply only unless processing (ie assumption) is not conducted specifically to treat someone differently based on the created inference.[30] As ICO argues, the key question here is not whether the inferences are correct, but whether a data controller is using an inference linked to sensitive data to influence its activities (in any way).[31] In other words, it is irrelevant if the inferred data that an incoming guest is likely to be COVID-19 positive on the basis of their origin of travel is correct – all that matters is that someone is using that inference to influence their activities.

3.2. Risk 1: Unfair Treatment Based on a Person's Assumed Health Status – Reasonable Business Practices Erroneously Applied

The inferences drawn about data subjects from the collected data might determine how they are viewed and evaluated – which poses a great risk.[32] Inferences of virus positivity or negativity could be used to make business decisions – would the guests (inaccurately) marked as COVID-19 positive be separated from other guests? Would they be allowed to enter the hotel restaurant only at specific times, or forbidden to enter the gym?

There is no doubt that, at least until the virus is contained, the private sector will have to adjust its daily business – looking at the accommodation sector, a limitation of the number of

[27] The population of Spain as of 29 April 2020, <available at https://www.worldometers.info/world-population/spain-population/ >, accessed on 29 April 2020.

[28] An educated guess may be defined as '*a guess that is made using judgment and a particular level of knowledge and is therefore more likely to be correct'*, definition from Cambridge dictionary, <available at *https://dictionary.cambridge.org/dictionary/english/educated-guess*>, accessed on 29 April 2020.

[29] ICO guide on special category of data, <available at https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/special-category-data/what-is-special-category-data/>.

[30] Ibid.

[31] Ibid.

[32] Wachter, Mittelstadt *CBLR 2018,* p 49. For an interesting discussion on which area of law is best suited for the protection of individual rights when it comes to profiling, see Mann, M. and Matzner, T., Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination, [2019], *Big Data & Society*, 2019(2), <available at https://doi.org/10.1177/2053951719895805>.

people who occupy a gym or eat at the same time seems reasonable to ensure social distance – as long as it is based on objective criteria. But making these decisions based on the 'inferred' health status, as it was illustrated above, infringes data protection rules as well as data subjects' right to be reasonably assessed. Under certain conditions, this might also be classified as profiling.[33]

3.3. Risk 2: Unlawful Processing – Diminishing Key GDPR Principles

The above observations do not mean that just holding an origin of travel would require sensitive data conditions. However, processing these data to 'indicate' the likelihood of COVID-19 infection (as exemplified above under 3.1.) to make business decisions, would be processing of sensitive data – thus prohibited. The arguable difficulty is not only the different treatment of data subjects – but the 'hidden' processing without respecting core GDPR principles.

The GDPR presents 'seven commandments' for lawful data processing.[34] The hereby exemplified inference clearly diminishes almost (if not) all of them. Lawfulness and fairness and the principle of accuracy have already been elaborated as being diminished – inexistence of appropriate legal basis for processing (lawfulness) and (unfair) different treatment of data subjects based on the COVID-19 status inferences (fairness), which are (conceivably inaccurate – accuracy) inference of someone's health status. The purpose limitation principle requires that data processing has a clearly defined purpose (eg processing of origin of travel for registration of tourists with the public register) and that such data cannot be reused for a purpose incompatible with the original one (eg for inferring if a guest is COVID-19 positive)[35]. The transparency principle (unsurprisingly, one of the biggest challenges when it comes to

[33] GDPR defines profiling as '*any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements*', Article 4(4). In its guidelines on profiling, Article 29 WP states that '*profiling can create special category data by inference from data which is not special category data in its own right but becomes so when combined with other data*'. Likewise, an interesting observation of whether automated processing must completely exclude any human involvement to constitute profiling, Article 29 WP argues that '*Article 4(4) GDPR refers to 'any form of automated processing' rather than 'solely' automated processing (referred to in Article 22). Profiling has to involve some form of automated processing – although human involvement does not necessarily take the activity out of the definition*.' see Article 29 WP, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, 17/EN [2016], p 6.

[34] The key principles of the GDPR are: 1. lawfulness, fairness, and transparency, 2. purpose limitation, 3. data minimisation, 4. accuracy, 5. storage limitation, 6. integrity and confidentiality, 7. accountability. Article 5 GDPR.

[35] GDPR requires that personal data are '*collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes; further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes*'. Article 5(1) b GDPR. Also see Basin D. and Debois S. and Hildebrandt T., On Purpose and by Necessity: Compliance Under the GDPR, [2018], also in: Meiklejohn S., Sako K. (eds) *Financial Cryptography and Data Security. FC* 2018, *Lecture Notes in Computer Science*, vol 10957. p 20-37.

'technological' inferences) would suffer the most – transparency requires that relevant data processing information is provided to data subjects in concise, transparent, intelligible and easily accessible way. [36] It is inconceivable that an organization would make exemplified data practices transparent, which would make data subjects unaware of inferences created about them. Is there a way of making these practices legitimate and proportionate – eg, if a hotel would make inference practices (using places of origin to assess the likelihood of infection and accordingly organizing business) transparent to guests – would that be GDPR-compliant? Even if one could argue that appropriate legal bases exist, that processing is necessary for the legitimate purpose and transparent, and that other GDPR principles are respected – it is confident to say that health status should not be subjected to guessing or assuming by a hotel, and only healthcare experts are authorized to inspect and conclude of someone's health.

4. Conclusion: The Future of Inferences and the Path to Clarity

Personal data inferences have been addressed as a significant loophole of the GDPR and this problem might become even more important during and post the COVID-19 crisis. Thus, this is a valuable opportunity to tackle this issue – both considering the crisis and for the future.

A 'unique' necessity test[37] on health data processing outside the healthcare industry seems to be a good start. Providing the private sector with clear guidance might encourage them to consistently ensure that the data processing is permissible under applicable legislation. The effectiveness of such a necessity test would be extremely beneficial if created within a dialogue between regulators and the industry – allowing both sides to be aware of which personal data might be necessary to contain the risk of COVID-19 and to ensure that the business is performed as normally and safely as possible, respecting the prohibition of sensitive data processing. This might also control (doubtfully accurate) inferences when it comes to different treatment during the COVID-19 crisis – at least, it would make industry more aware of the personal data they process. The assumptions and determinations of a person's COVID-19 status should stay

[36] For the analysis of the transparency requirements, see Article 29 WP, Guidelines on transparency under Regulation 2016/679, 17/EN, [revised 2018], <available at https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227>. For additional observations of the compliance with transparency principles in practice, see Mohan J. and Wasserman M. and Chidambaram V., Analysing GDPR Compliance Through the Lens of Privacy Policy, [2019], also in: Gadepally V. et al. (eds) *Heterogeneous Data Management, Polystores, and Analytics for Healthcare. DMAH 2019*, Poly 2019. *Lecture Notes in Computer Science*, vol 11721. Springer, Cham, p 82-95.

[37] Necessity test in its essence would mean assessing whether the processing (of a certain category of data in question) is necessary for a clearly defined purpose. See, for example, ICO Guide to the General Data Protection Regulation (GDPR), p 82, available via <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/lawful-basis-for-processing/legitimate-interests/>.

exclusively in the hands of healthcare institutions and should be shared with the private sector (provided that personal data is being shared) strictly in line with data protection principles.

It seems that sometimes we get so 'trapped' in the technology surrounding us that we forget that data manipulation existed much before its advancement. Learning how to handle 'simple' inferences might get us one step closer to finding a solution for more sophisticated technology-created ones. Thus, looking beyond COVID-19, EU legislation would need an update to deal with the threat of inferences.[38] The priority of regulators should be to adopt precise guidelines dealing exclusively with inferred data processing. Although it would be quite a challenging task to predict and regulate all practical examples of personal data under GDPR, the legal treatment of inferences should be clearly drawn. This would potentially improve the overall data protection scheme, provide the private sector with a clear framework for processing individuals' data and create an environment which supports and invigorates data subject's rights.

[38] The controversial and long-awaited E-privacy Regulation might provide for additional regulation and clarification of inferences as well, see Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications), COM/2017/010 final-2017/03 (COD).

Bibliography

Secondary sources

Basin D, Debois S and Hildebrandt T, On Purpose and by Necessity: Compliance Under the GDPR, [2018], also in: Meiklejohn S, Sako K (eds) Financial Cryptography and Data Security. FC, Lecture Notes in Computer Science, vol 10957, 2018, p 20-37

Dzięgielewska, O, Anonymization, tokenization, encryption. How to recover unrecoverable data, [2017], Computer Science and Mathematical Modelling, 2017(6): p 9-13

Etteldorf, C, COVID-19 Special EU Member State Data Protection Authorities Deal with COVID-19: An Overview, [2020], European Data Protection Law Review, 2020(2) preview <available at: https://www.lexxion.eu/wp-content/uploads/2020/03/COVID-19-Special-Data-Protection-Authorities-Deal-with-COVID-19.pdf>

Fischer-Hübner, S and Angulo, J and Karegar, F and Pulls, T, Transparency, Privacy and Trust – Technology for Tracking and Controlling My Data Disclosures: Does This Work? [2016], also published in: Habib S, Vassileva J, Mauw S, Mühlhäuser M (eds) (Trust Management X. IFIPTM), IFIP Advances in Information and Communication Technology, 2016 (473), p 3-14

Hu, R, Stalla-Bourdillon, S, Yang, M, Schiavo, V and Sassone, V, Bridging Policy, Regulation, and Practice? A Techno-Legal Analysis of Three Types of Data in the GDPR, also published in Data Protection and Privacy: The Age of Intelligent Machines [2017], Leenes Rosamunde van Brakel, R, Gutwirth, S and De Hert, P (eds) (Hart Publishing, 2017) <available at SSRN https://ssrn.com/abstract=3034261>

Malgieri, G.and Comandé, G. Sensitive-By-Distance: Quasi-Health Data in the Algorithmic Era, [2017], Information, Communication and Technology Law, 2017 (3), <available at SSRN https://ssrn.com/abstract=3020628>

Mann, M.and Matzner, T. Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination, [2019], Big Data & Society, 2019(2), <available at https://doi.org/10.1177/2053951719895805>

Mohan J, Wasserman M and Chidambaram V, Analyzing GDPR Compliance Through the Lens of Privacy Policy, [2019], also in: Gadepally V et al (eds) Heterogeneous Data Management, Polystores, and Analytics for Healthcare. DMAH 2019, Poly 2019. Lecture Notes in Computer Science, vol 11721. Springer, Cham, pp 82-95

Nadezhda, P, The law of everything. Broad concept of personal data and future of EU data protection law, [2018], Law, Innovation and Technology, 10:1, p 40-81, <available at: https://doi.org/10.1080/17579961.2018.1452176>

Pandit HJ, Fernández JD, Debruyne C and Polleres A, Towards Cataloguing Potential Derivations of Personal Data, [2019], also in: Hitzler P et al (eds) The Semantic Web: ESWC 2019 Satellite Events. ESWC, Lecture Notes in Computer Science, 2019(11762), p 147-151

Ross, D, The Economic Impact of COVID-19 on the Hotel Industry , Hospitality Tecnology, [2020], <available at: https://hospitalitytech.com/economic-impact-covid-19-hotel-industry>

Wachter, S and Mittelstadt, B, A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI [2018], Columbia Business Law Review 2019(2) p 22 <available at SSRN:https://ssrn.com/abstract=3248829>.

Wachter, S, Normative challenges of identification in the Internet of Things: Privacy, profiling, discrimination, and the GDPR, [2018], Computer Law and Security Review 2018(3), <available at: https://doi.org/10.1016/j.clsr.2018.02.002>

Vedder, A, Why Data Protection and Transparency Are Not Enough When Facing Social Problems of Machine Learning in a Big Data Context, [2018], <available at SSRN:https://ssrn.com/abstract=3407853>, also in Emre Bayamlioglu et al. (eds), Being profiled: Cogitas, ergo sum. 10 Years of Profiling the European Citizen. Amsterdam University Press, 2018.

Case law

Joined Cases C–141 & 372/12, *YS, M and S v. Minister voor Immigratie, Integratie en Asiel*, [2014] E.C.R. I-2081

Case C-434/16, *Peter Nowak v. Data Prot. Comm'r*, [2017] E.C.R. I994

Law and policy documents

Article 29 WP, Guidelines on the Right to Data Portability, 16/EN, p 9–11 [2016], <available at https://ec.europa.eu/newsroom/document.cfm?doc_id=44099>

Article 29 WP, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, 17/EN [last revised 2018]

Article 29 WP, Guidelines on transparency under Regulation 2016/679, 17/EN, [revised 2018], <available at https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227>

Annex to the letter of Article 29 WP responding to a request of the European Commission to clarify the scope of the definition of health data in connection with lifestyle and wellbeing apps, [5 February 2015], <available at https://ec.europa.eu/justice/article29/documentation/otherdocument/files/2015/20150205_letter_art29wp_ec_health_data_after_plenary_annex_en.pdf>

Belgian Data Protection Authority guidlines for employees' data processing during COVID-19, <available at: https://www.autoriteprotectiondonnees.be/covid-19-et-traitement-de-donn%C3%A9es-%C3%A0-caract%C3%A8re-personnel-sur-le-lieu-de-travail>

Commissioner's Office (ICO) guide on special category of data, <available at https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/special-category-data/what-is-special-category-data/>

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free

movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), published in Official Journal of the European Union, L 119, [4 May 2016]

Other sources

Operational considerations for COVID-19 management in the accommodation sector, Interim guidance, [31 March 2020], <available at: https://apps.who.int/iris/bitstream/handle/10665/331638/WHO-2019-nCoV-Hotels-2020.1-eng.pdf>

Statement by the EDPB Chair on the processing of personal data in the context of the COVID-19 outbreak as of 16 March 2020, <available at: https://edpb.europa.eu/news/news/2020/statement-edpb-chair-processing-personal-dana context-covid-19-outbreak_en>

https://iapp.org/resources/article/dpa-guidance-on-covid-19/

https://www.enforcementtracker.com/

/http://paulvines.com/2018/02/21/gdpr-and-latent-variables.html