FACULTY OF LAWS

Dr Michael Veale
Associate Professor in Digital Rights and Regulation
e-mail: m.veale@ucl.ac.uk

11 March 2022

European Data Protection Board

Via e-mail

To whom it concerns,

Consultation Response on Guidelines 01/2022 on data subject rights - Right of access

Thank you for preparing the draft *Guidelines 01/2022 on data subject rights - Right of access*. I write as an academic who attended the EDPB workshop on these guidelines in 2019, and subsequently co-authored a piece to inform the drafting of these guidelines, "Getting Data Subject Rights Right", published in JIPITEC.[1] Below please find some feedback on specific aspects of these Draft Guidelines.

Access Rights and Unstructured Data

The Guidelines should be clearer that some data is unstructured, yet access rights still apply. It currently states (p. 3):

> *The controller will have to search for personal data throughout all IT systems and non-IT filing systems based on search criteria that mirrors the way in which the information is structured, for example name and customer number.*

However, some personal data, particularly data such as text and images, are unstructured. For example, the recent action by DPAs including the Italian and French authorities regarding the data held by *Clearview AI*, which is stored and searched by facial template and may not be structured as a conventional database. More frequently, data forms free-text in emails or documents, and is searched by reference to free-text terms across email and similar document management systems. Consequently, the **guidelines should acknowledge that some data are not structured, but still within scope, and search queries should be adapted as to the form of the data**. In particular, it is worth noting that there is no "filing system" limitation on the right of access — all automated processing is within scope of data protection, and therefore protected by this right.

Data format

Paragraph 32 states that an example of a commonly-used electronic form is a PDF. However, it is worth noting that PDFs are formats that make data difficult to analyse, and are often applied unnecessarily by data controllers. While there is no obligation to transmit data in a machine-readable form, unlike as is the case for Article 20 GDPR, I am concerned, and have experienced, data controllers applying the PDF file-format to data under Article 15 unnecessarily, making it difficult to access or analyse. This is particularly important as data held by controllers grows in the digital economy, and data subjects may find themselves with have the increasing ability (through technologies such as AI) to scrutinise even complex data themselves. Where data are only available in PDF format, this is not a problem. But where data are available in an eminently machine readable format, such as JSON or CSV, which are easy to open for users as they are just text files, data controllers should not, pursuant to the fairness principle, be allowed to *degrade* the quality of the data, which would *only serve to disadvantage the data subject* and abuse the asymmetric power

[1] Jef Ausloos and others, 'Getting Data Subject Rights Right: A Submission to the European Data Protection Board from International Data Rights Academics, to Inform Regulatory Guidance' (2020) 10 JIPITEC.

Faculty of Laws, University College London
Bentham House, Endsleigh Gardens, WC1H 0EG
Tel: +44 (0)2031 089736
m.veale@ucl.ac.uk

relations, and impede their analysis of data received. PDF files are notorious for introducing errors (especially spacing) and making tables impossible to export, because they are designed for printers, not for computing of the information within them. **The EDPB should state that the data should be provided in the most machine-readable format that available, and while controllers are not obliged to render data machine readable where it is not or would not be used or understood this way generally (e.g. a scanned document, or a free-text email), they are also obliged not to render it *less* machine-readable where machine readability is crucial to understand it (e.g. tabular data, numeric data, sensor observations, statistical inferences).**[2]

Advertiser IDs and similar.

Data subjects face difficulties accessing data that is catalogued by identifiers that are hidden inside web browsers or operating systems, particularly concerning cookies and similar technologies.[3] The Belgian DPA has identified the online advertising environment as one where data protection by design places an obligation on joint controllers to facilitate access to data rights, which may involve facilitating access to the provision of identifiers.[4]

Paragraph 67 of the Draft Guidelines provides some welcome guidance but does not quite clarify this far enough. In particular, while it notes that the "additional information" will be the cookie identifier, it should highlight that the principles of data protection by design and by default oblige the controller to have designed the system to assist the user in accessing this information in order to secure their right of access, especially where this is in practice *necessary* to secure the rights. Without this, the right is a dead letter as the controller can deliberately engineer a system to make it difficult for the user to provide this "additional information" when this difficulty is *entirely artificial*. **The EDPB should indicate that data protection by design and by default would apply in obliging controllers to help secure access to "additional information" such as cookie identifiers, where the barriers are primarily that they are designed to be difficult to access, and where they are required for access rights to be secured**.

Ex post information on recipients and automated decisions.

Paragraphs 115 and 118 of the Draft Guidelines, stating that recipients should be provided in an Article 15 request even where they were unknown or not specifically tailored to a user in an Article 13–14 transparency obligation, is to be welcomed. **The EDPB should retain its guidance on tailored and ex post clarification of particular recipients via Article 15, once they are known to the controller**.

Similarly, paragraph 119 on ex post information about automated decision-making is welcome, and in line with academic research in the area.[5] **Paragraphs 115, 118 and 119 should be retained**.

Residual Data from Self-Service Access Systems

---

[2] For further on this, see ibid paras 23–25.

[3] Michael Veale and others, 'When Data Protection by Design and Data Subject Rights Clash' (2018) 8 International Data Privacy Law 105; Chris Norval and others, 'RECLAIMING Data: Overcoming App Identification Barriers for Exercising Data Protection Rights' in *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers* (UbiComp '18, New York, NY, USA, ACM 2018).

[4] Belgian Data Protection Authority, 'Decision on the Merits 21/2022 of 2 February 2022, Complaint Relating to Transparency & Consent Framework (IAB Europe), DOS-2019-01377' (2 February 2022) <https://www.autoriteprotectiondonnees.be/publications/decision-quant-au-fond-n-21-2022-english.pdf> accessed 2 February 2022.

[5] Andrew D Selbst and Julia Powles, 'Meaningful Information and the Right to Explanation' (2017) 7 International Data Privacy Law 233.

Faculty of Laws, University College London
Bentham House, Endsleigh Gardens, WC1H 0EG
Tel: +44 (0)2031 089736
m.veale@ucl.ac.uk

Paragraph 136, clarifying that self-service systems should never limit the information received and do not prejudice a written request, should integrate the principle of fairness. Many systems do not provide all of the data that a controller may hold on an individual. For example, Google Takeout would not provide internal company emails relating to that user. Some data may require additional information, like cookie IDs (see above). **The Draft Guidelines should state that where self-service tools are provided, the principle of fairness obliges them to inform the user if they are only providing a subset of all the information users have a right to, and direct them to the Article 13/14 information (the privacy policy) where they can contact to access the remainder.** Without this, data subjects may be confused and may think that the totality of the right of access is whatever the controller has designed it to appear.

Making log files clear

The Draft Guidelines in paragraph 138 states that log files are an example of a file that needs to be made clear by controllers. This is welcome. **However, the EDPB should give an example of what "careful and thorough" presentation is in this context.**

Trade secrets

Paragraph 171, example 2, concerning trade secrets and stocking measurements, is easily able to be abused by data controllers and should be removed. This opens the door to both simplistic measurements and invasive profiling techniques being denied to those accessing data about them on the spurious claim of trade secrets, which are impossible for data subjects to verify and difficult for DPAs to as well. In practice, controllers will seek to claim that even the number of variables, or the categorisation approach, or the label applied to users, are revealing of the upstream analytic method, which they will claim is a trade secret. This is balancing better left to the CJEU, should it emerge, rather than placed in this guidance, where it will significantly weaken the right to access in a digital economy beyond the legitimate protection of genuine trade secrets. **The EDPB should remove example 2 in paragraph 171**.

Access to backups "where technically feasible"

The EDPB state in Paragraph 108 that archived data should only be provided access to "where technically feasible". This is not clarified in the text, and is not a valid wording of an exemption in the GDPR. **"[W]here technically feasible" should be replaced with "unless technically impossible".**

I am at your disposal for any further questions on these areas.

Yours sincerely,

Dr Michael Veale

Faculty of Laws, University College London
Bentham House, Endsleigh Gardens, WC1H 0EG
Tel: +44 (0)2031 089736
m.veale@ucl.ac.uk