Nobody puts data in a corner? Why a new approach to categorising personal data is required for the obligation to inform

Emma Cradock·*, David Millard*, Sophie Stalla-Bourdillon**

** Electronics and Computer Science, University of Southampton, Southampton, UK*
*** School of Law, University of Southampton, Southampton, UK*

ABSTRACT
Transparency is a key principle of EU data protection law and the obligation to inform is key to ensuring transparency. The purpose of this obligation is to provide data subjects with information that allows them to assess the compliance and trustworthiness of the data controller. Despite the benefits of categorising personal data for this purpose, a coherent and consistent approach to doing so under the obligation to inform has not emerged. It is unclear what a 'category' of personal data is and when this information must be provided. This results in reduced transparency for data subjects and uncertainty for data controllers regarding their legal obligations, defeating the purpose of this obligation. This article highlights these issues and calls for clarification on them. It also posits that in clarifying the law, a new approach to categorising personal data is required, to achieve the benefits of categorisation and increase the transparency of personal data processing for data subjects.

Keywords: data protection, transparency, personal data, categories, obligation to inform

1. Transparency and the obligation to inform

'Transparency' has always been a key principle of the European Union (EU) data protection framework, but its importance has been made explicit under the new General Data Protection Regulation (GDPR)[1]. The GDPR replaces Directive 95/46/EC (DPD)[2] as the main instrument of data protection regulation within the EU. In particular, the increased importance of transparency is signified by the introduction of the words *'and in a transparent manner'* to the end of the first data protection principle. Previously, this principle simply stated that personal data must be processed *'fairly and lawfully'*[3].

· Electronics and Computer Science, University of Southampton, Building 32, Highfield Campus, Southampton, UK, SO17 1BJ
Email address: erc1e10@soton.ac.uk

[1] 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) <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.119.01.0001.01.ENG&toc=OJ:L:2016:119:TOC> accessed 29 May 2016

[2] Directive 95/46/EC of the European Parliament and of the Council of on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281/31

[3] See Article 5(1)(a) GDPR, which will replace Article 6(1)(a) DPD.

Whilst there is no agreed definition of 'transparency', it is generally the idea that data controllers should keep data subjects informed of how their personal data is being (and will be) used[4]. This information then enables data subjects to identify compliant organisations that can be trusted with their personal data[5]. Importantly, transparency is not simply desirable, it is vital to the efficacy of the EU data protection framework as a whole. As a rights-based, complaint-driven system[6], the framework's success is reliant upon data subjects enforcing their rights and keeping a check on data controllers. Transparency is vital in enabling individuals to do this.

Under the framework, transparency is initially created through the data controller's 'obligation to inform'[7], sometimes referred to as the 'right to information'. This requires controllers to inform data subjects of certain information upon obtaining their data. Transparency is then continually provided for through the data subject's right of access to their data[8]. A certain amount of transparency is also created under the current data controller obligation to notify the supervisory authority[9]. However, this only concerns the data controller's processing activities as a whole and will be replaced with an obligation to internally document this information instead under the GDPR[10].

Thus, of these, the obligation to inform is especially key to making data processing transparent. It provides the only information data subjects are guaranteed to receive about processing with no further effort on their part (unlike enforcing their right of access). Furthermore, information is generally given at the time the personal data is obtained. This ensures that data subjects have the information required to make informed and appropriate decisions where they have a choice over processing[11]. Where there is no choice, the information helps data subjects: understand what is happening with their personal data; enforce their data protection rights (when necessary)[12]; and detect any unlawful, or questionable practices.

At the inception of the DPD, the majority of the personal data processed by data controllers was 'provided' by individuals, with their full awareness that their personal data was being obtained[13]. As the active source, the presumption was that both data subjects and controllers had equal information on exactly what **'**personal data' was being collected and processed. However, technological progress over the last twenty years has seen a substantial growth in the amount of personal data that is

---

[4] European Union Agency for Fundamental Rights, Council of Europe and the Registry of the European Court of Human Rights 'Handbook on European data protection law' <http://www.echr.coe.int/Documents/Handbook_data_protection_ENG.pdf> accessed 29 May 2016

[5] Information Commissioner's Office, 'Overview of the General Data Protection Regulation (GDPR)' <https://ico.org.uk/media/for-organisations/data-protection-reform/overview-of-the-gdpr-1-0.pdf> accessed 29 September 2016

[6] Mistale Taylor 'Safeguarding the Right to Data Protection in the EU, 30th and 31st October 2014, Paris, France' [2015] Utrecht J. Int'l & Eur. L., 31, 145.

[7] See Articles 13 and 14 of the GDPR, previously Articles 10 and 11 of the DPD.

[8] See Article 15 of the GDPR, previously Article 12 of the DPD.

[9] See Articles 18 and 19 DPD.

[10] See Article 30 GDPR concerning Records of Processing Activities.

[11] Perri 6, *The future of privacy: Volume 1 Private life and public policy* (Demos 1998)

[12] BEUC 'A Comprehensive Approach on Personal Data Protection in the European Union, European Commission's Communication: The European Consumers' Organisation's response' <http://ec.europa.eu/justice/news/consulting_public/0006/contributions/organisations/b euc_en.pdf> accessed 20 April 2016

[13] OECD Working Party On Security And Privacy In The Digital Economy 'Protecting Privacy in a Data-driven Economy: Taking Stock of Current Thinking' <http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg%282014%293&doclanguage=en> accessed 29 May 2016

observed, derived and inferred, without the awareness of the individual[14]. At the same time, the definition 'personal data' has broadened, to account for new technologies[15]. Data processing tools have become increasingly powerful, sophisticated, ubiquitous, and inexpensive, making information easily searchable, linkable and traceable[16]. The result of this progress is that it can no longer be presumed that data subjects are aware of what 'personal data' is being collected or how it can be generated and processed by data controllers.

In its role of redressing the balance of information between data subjects and data controllers, it is the task of the 'obligation to inform' to reverse this presumption, and require that data controllers provide information to individuals that makes what personal data is being collected and processed transparent. If not, data processing will be less transparent now than it was twenty years ago, as data subjects will have access to less information about the processing of their data, and will be less capable of assessing the compliance of controllers.

In theory, appropriately categorising personal data and informing individuals of the 'categories' being processed could provide the first step in filling this lacuna. Yet, a coherent and consistent approach to doing so under the obligation to inform has not emerged. This article begins by highlighting the benefits of categorising personal data. It then highlights and the uncertainty in the law on both what a 'category' of personal data is in relation to the obligation to inform, and when a data subject must be informed of these. In calling for clarification, the article examines the current approaches to categorisation and concludes that a new approach to categorising personal data is required, through which meaningful information that increases the transparency of data processing can be provided. Whilst proposing and describing a new approach is beyond the scope of this article, it provides a discussion of the benefits to be achieved and issues to be avoided, against which any proposal of a new approach should be assessed.

2 The benefits of categorising personal data

A 'category' is defined as a *'class of people or things with shared characteristics'*[17]. The purpose of creating categories in relation to any phenomena is to reduce the number of discriminations in the world, so that each individual thing does not have a separate label[18]. Categorisation allows an individual to ascertain information about 'things', simply from knowing to which category they belong[19].

Given the purposes of categorisation in general, in theory, an **appropriate categorisation of personal data** could provide a number of benefits, increasing the transparency of personal data

[14] As defined by the OECD Expert Roundtable (see supra), 'Observed' data is personal data that is observed by others and recorded in a digital format e.g. cookie data or sensor data. 'Derived data' are data generated from other data, after which they become new data elements related to a particular individual e.g. a calculation of customer profitability based on the ration between number of visits and items bought. 'Inferred Data' are the product of probability-based analytic processes. They are the result of the detection of correlations, which are used to create predictions of behaviour e.g. the likelihood of future health outcomes based on an analysis of large and diverse medical data sets.

[15] Gerrit-Jan Zwenne, Diluted Privacy Law (April 12, 2013). Available at SSRN: https://ssrn.com/abstract=2488486

[16] OECD, 'The Evolving Privacy Landscape: 30 Years After the OECD Privacy Guidelines' <http://www.oecdilibrary.org/docserver/download/5kgf09z90c31.pdf?expires=1477319599&id=id&accname=guest&checksum=3CCA414864E0A35F15FA336CB015E5B8> accessed 20 October 2016

[17] *Oxford English Mini Dictionary* (First published 1981, Oxford University Press, 2011) 82

[18] Eleanor Rosch and Barbara Lloyd, (eds) *Cognition and categorization* 27-48. (1978, Lawrence Erlbaum)

[19] Eleanor Rosch and Barbara Lloyd, (eds) *Cognition and categorization* 27-48. (1978, Lawrence Erlbaum)

processing. Some benefits would come from the information gained simply from knowing which category of personal data is processed, and some come from using the category as an anchor, to which further information about processing can be attached.

2.1 Benefits from knowing the category

Appropriately categorising personal data and knowing the category can:

- ***Enable an assessment of the risk involved in the processing.*** Understanding the differences between categories of personal data allows for an assessment of the different risks involved in their processing. For data controllers, this is helpful when conducting a Data Protection Impact Assessment (Article 35 GDPR). For supervisory authorities it informs their decisions on the amount that an administrative fine should be, or the extent that enforcement measures should take. For example, Article 83(2)(g) GDPR states that *'when deciding whether to impose an administrative fine and deciding on the amount of the administrative fine in each individual case, due regard shall be given to … the categories of personal data' affected by the infringement'.* Thus, appropriately categorising personal data and understanding the differences between categories processed can inform decisions on the severity of the risk and therefore, punishment required. It can also inform assessments of the adequacy of the level of protection afforded by a third country. Article 25(2) DPD, states that the *'adequacy of the level of protection afforded by a third country shall be assessed in the light of all the circumstances surrounding a data transfer… particular consideration shall be given to the nature of the data'.* Thus, understanding how the personal data processed differs from other personal data can enable this assessment.

- ***Inform an assessment of the appropriate technical and organisational measures that should be in place to ensure security.*** Following on from understanding the risk, knowing the categories of personal data that are processed and the differences between them (including the risk) can inform decisions on what technical and organisational measures are necessary to ensure the security of personal data. Recital 46 and Article 17(1) DPD confirm that the *'nature of the personal data to be protected'* should be taken into account when making sure that technical and organisational measures ensure the appropriate level of security. Such an assessment must also take into account the state of the art, and the costs of their implementation in relation to the risks inherent in the processing. Therefore, an organisation must understand what different categories of personal data it processes as a whole, in order to understand what the appropriate technical and organisational measures will be and whether different levels of this are involved.

- ***Aid decisions on whether a secondary purpose is compatible.*** A key principle of the current and future EU data protection framework is purpose limitation (Article 5(1)(b) GDPR and Article 6(1)(b) DPD). The concept of purpose limitation has two parts, that personal data must be collected for *'specified, explicit and legitimate'* purposes (purpose specification) and that it not be *'further processed in a way incompatible'* with those purposes (compatible use)[20]. Recital 50 GDPR states that understanding the *'nature of the personal data'* helps ascertain whether a further purpose is compatible. Article 6(4)(c) GDPR elaborates on this,

[20] Article 29 Data Protection Working Party 'Opinion 03/2013 on purpose limitation' <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf> accessed 20 October 2016

and confirms that (where consent or a Union or Member State law does not apply) consideration of whether a purpose is compatible with the purpose for which the data were originally collected, should take into account *'the nature of the personal data, in particular whether special categories of personal data are processed, pursuant to Article 9, or whether personal data related to criminal convictions and offences are processed, pursuant to Article 10'*. Therefore, understanding the 'nature' of the personal data, and which categories are processed, can support an assessment of whether a purpose is compatible.

- ***Reduces the amount of information that needs to be provided to increase transparency.*** If personal data were appropriately divided into categories that allowed individuals to understand more about the personal data processing simply from knowing which category (or categories) were processed, this could reduce the amount of information currently required to create this understanding. The benefit here is similar to the benefit envisaged by providing standardised icons under Article 12(7) GDPR i.e. giving a meaningful overview of the processing. One of the key criticisms of the manifestation of the obligation to inform is that it generally results in long and complicated privacy notices, which are never read[21]. This reduces the transparency of processing because by not reading privacy policies, in practice, data subjects understand very little information about the processing of their personal data. Thus, appropriately categorising personal data has the potential to reduce the amount of information that needs to be provided, removing a disincentive for engaging with this information.

2.2 Benefits from using categories as anchors for further information

There are also benefits of categorisation that can be realised by appropriately categorising personal data and then using the categories as an anchor, to which further information can be attached. Further information that can be provided includes:

- ***Further information about processing.*** For example, Article 30(1)(f) of the GDPR requires that a data controller (in keeping a record of its processing activities) record *'where possible, the envisaged time limits for erasure of the different categories of data'*. Thus, first specifying the categories of personal data processed then allows other information such as 'time limits for erasure' to be attached to them. This is also true for other information, such as sources of categories of personal data. This information provides a more granular view of the processing.

- ***The identification of responsibilities in relation to different categories***. For example, Article 28(3) GDPR mandates that processing by a processor must be governed by a contract, or other binding legal act under Union or Member State law, which (amongst other information) must set out the *'type of data'* to be processed. Therefore, describing and differentiating between personal data facilitates an understanding of which personal data a processor is responsible for. This allows the processor to know what to document, but also for a supervisory authority or court to understand where responsibility lies. Appropriate categorisation could be used for this purpose.

- ***The attachment of different levels of protection or obligations and rights, in relation to different personal data.*** For example, under both the DPD (Article 8(1)) and GDPR (Article

[21] Aleecia M McDonald and Lorrie Faith Cranor, 'The cost of reading privacy policies' [2008] *ISJLP* 4: 543

9(1)), 'special categories of personal data' are defined. Article 9(1) GDPR defines these categories as the *'Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation*. These categories are then given a higher level of protection than 'normal' personal data, in that processing of them is prohibited, unless under an applicable exception. Thus, by identifying categories, a different level of protection can be attached, including different rights and obligations. This has the benefit of providing a more nuanced approach to data protection. This could be used to support a risk-based approach to regulation, which is arguably required giving the ever-expanding definition of personal data[22].

2.3 Summary

As can be seen there are many benefits to be achieved by appropriately categorising personal data. It can enable an assessment of the risk involved in the processing; dictate appropriate technical and organisational measures; and inform decisions on whether a secondary purpose is compatible. Categories can also be used to provide a more detailed description of processing by allowing further information such as storage periods to be attached to them or the stakeholder responsible for processing them. Where different rights or obligations apply to different categories it also allows the data subject to assess compliance of the data controller in light of their applicable obligations when they are informed of the categories processed.

Given these benefits, and the goal of transparency being to enable data subjects to identify organisations that are compliant who can be trusted with their personal data[23], it would seem logical that there would be a consistent and robust approach to categorising personal data under the obligation to inform. It would also seem logical that there would be a requirement that the subject must always be informed of the category or categories of personal data being processed. This last point is especially so, as other stakeholders such as data controllers and supervisory authorities are seen to need (and be entitled to) this information. Yet, in practice, this is not the case. It is unclear what a 'category' of personal data is under the obligation to inform, and as the next section discusses, the framework is inconsistent on when data controllers must inform individuals of the categories of personal data they process under the obligation.

3 When should a data controller provide the 'categories of personal data' under The Obligation to Inform?

3.1 The Data Protection Directive

The 'obligation to inform' is currently provided for in Articles 10 and 11 of the DPD. Article 10 governs cases of collection **from the data subject,** and Article 11 governs cases **where the data is not obtained from the data subject**. Both state that data subjects should be provided with at least:

- The identity of the controller and his representative, if any;

---

[22] Gerrit-Jan Zwenne, Diluted Privacy Law (April 12, 2013). Available at SSRN: https://ssrn.com/abstract=2488486

[23] Information Commissioner's Office, 'Overview of the General Data Protection Regulation (GDPR)' <https://ico.org.uk/media/for-organisations/data-protection-reform/overview-of-the-gdpr-1-0.pdf> accessed 29 September 2016

- The purposes of the processing for which the data are intended; and
- Any further information necessary, having regard to the specific circumstances, to guarantee fair processing in respect of the data subject.

The first two points make it relatively clear what information must be provided, but the last point is a wide and case-specific requirement. To provide further clarity, both Article 10 and 11 DPD each provide three examples of information that might fall within this third point, and could be necessary to inform data subjects of (depending on the circumstances). Both Articles provide the examples of informing data subjects:

- The recipients or categories of recipients;
- The existence of the right of access to and the right to rectify the data concerning him

However, the Articles differ on the third example. Article 10(c) DPD (governing collection **from** the data subject) provides the example of informing the data subject of *'whether replies to the questions are obligatory or voluntary, as well as the possible consequences of failure to reply'*. Whereas, Article 11(c) DPD includes an example that when information is obtained 'not from the data subject', they **may** need to be informed of the *'the categories of data concerned'*.

The fact that informing data subjects of the 'categories of data concerned' is not stated anywhere within the text of Article 10 DPD can arguably be interpreted as meaning that whenever personal data is obtained from the data subject, the data controller does not have to inform them of the categories of data being obtained. The lack of this example under Article 10(c) DPD compared to its inclusion under Article 11(c) DPD would support this interpretation.

However, although it is not listed as an example, it could still be legally required, under Article 10(c) DPD. This is because the examples provided are not exhaustive, and technically Article 10(c) DPD requires the data controller to inform the individual of **any information required** for the processing to be 'fair'. Thus, it could be interpreted that in certain circumstances (e.g. where the individual is not fully aware of the data being collected) a data controller **would** be obligated to inform them of the 'categories of data concerned', for the processing to be 'fair'.

Yet, even if this were so, a case-by-case necessity test is taken to any information deemed necessary under this section. This was confirmed by the European Commission in their first report on the implementation of the DPD[24]. Thus, although the argument could be made that a data controller **could** be obligated under this section (in certain circumstances) to inform the data subject of the 'categories of data concerned', there is currently no binding legal precedent stating that data subjects **must** be informed of the categories of personal data obtained from them. Moreover, even if there were, it would depend on the circumstances in which it was deemed necessary as to how far this obligation would extend. If such a ruling were made under Article 10(c), then a data controller would only be deemed to be under an obligation in circumstances similar to those in the ruling.

Whereas, informing individuals of the 'categories of data concerned' **is** mentioned within Article 11 DPD (governing data collection not from the data subject), but only as an **example** of 'any further information necessary' under Article 11(c) DPD. This means that currently, it is not mandatory for

[24] European Commission, 'First report on the implementation of the Data Protection Directive (95/46/EC)' <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52003DC0265&from=EN> accessed 9 November 2015

data controllers to inform data subjects of this in every case under Article 11 DPD, only when it is necessary to guarantee 'fair' processing. Compared to Article 10 DPD, this at least indicates that the DPD envisions that there **are** situations in which data controllers will be obligated to provide individuals with the 'categories of data concerned' when personal data is not obtained from the data subject. However, it does little to clarify what these circumstances will be.

3.1.1 Further uncertainty under the Data Protection Directive

Further uncertainty arises under the DPD when considering whether it is Article 10 or Article 11 DPD that applies. The obligation to inform distinguishes between situations where the data is 'collected from the data subject' (Article 10 DPD) and where the data is 'obtained not from the data subject' (Article 11 DPD). However, looking at the differences between 'provided', 'observed', 'inferred' and 'derived' personal data in relation to the obligation to inform (See n 14), it is unclear which situation applies where.

It seems clear that 'provided' data would fall under Article 10 DPD, because it is provided with the awareness of the data subject (and therefore certainly obtained from them)[25]. Yet, for personal data that is 'observed' by others and recorded in a digital format[26] e.g. data originating from online cookies or sensors, how would this be classed? It could be argued that the individual is the source of the data, as their actions generate the data in some way and therefore Article 10 DPD would apply. However, it could also be argued that the cookie or the sensor is the source of the data and therefore Article 11 DPD would apply. This confusion is also true for data that is 'inferred' or 'derived'. In both cases, data is generated from other data. It is unclear whether the obligation to inform only covers the instance of the original personal data collection (i.e. the data which is then used to generate other personal data, such as items bought and number of visits) or whether as the data controller begins deriving and inferring personal data from this (e.g. the profitability of the individual), that they are under an Article 11 DPD obligation, because this new personal data that is being created has not strictly been obtained from the individual. For example, this could be an instance where a data subject should be informed of the 'categories' of personal data processed under Article 11 DPD.

Therefore, it is not always clear which Article the data controller's processing activities are governed by, making it unclear exactly what their obligations are. This distinction becomes even more important where the information requirements between the two Articles differ, especially under the GDPR as discussed in Section 3.4.

3.2 The Article 29 Working Party Guidance

Although from the hard law of the DPD it is currently unclear when, and whether, data controllers need to inform individuals of the 'categories of data concerned', the Article 29 Working Party[27] (WP) has repeatedly referred to a need for data controllers to be informing individuals of the

---

[25] OECD Working Party On Security And Privacy In The Digital Economy 'Protecting Privacy in a Data-driven Economy: Taking Stock of Current Thinking' <http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg%282014%293&doclanguage=en> accessed 29 May 2016

[26] OECD Working Party On Security And Privacy In The Digital Economy 'Protecting Privacy in a Data-driven Economy: Taking Stock of Current Thinking' <http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg%282014%293&doclanguage=en> accessed 29 May 2016

[27] Article 29 Data Protection Working Party is an independent body that gives expert advice on data protection within the EU.

personal data they process under the obligation to inform. However, they have not always referred to the obligation using the term 'categories'.

Even as early as 1999, the WP were concerned about processing operations performed without an individual's knowledge, stating that *'Internet software and hardware products should provide the Internet users* ***information about the data that they intend to collect, store or transmit****'*[28]. The WP echoed this guidance again in relation to online data protection in 2000[29].

The WP have also stated that:

• Individuals should be given *'accurate and full information of all relevant issues, in particular those specified in Articles 10 and 11 of the Directive, such as* ***the nature of the data processed*** *'*[30];

• *'According to Article 10 ... each data subject has a right to know ... in the context of apps ...* ***what type of personal data is being processed****'* and that *'the relevant data controller must inform potential users at the minimum about: ...* ***the precise categories of personal data the app developer will collect and process****'*[31]

Although the latter was in the context of apps and smart devices, and the former in the context of electronic health records, for both, the WP based their opinion on Article 10 DPD, despite the lack of this requirement within this Article. If the WP based this guidance on Article 10(c) DPD, then this would not be an information requirement in every case, as the European Commission opined[32]. It could only be seen as a requirement in the specific circumstances referred to by the guidance (as discussed in Section 3.1).

Interestingly, in 2014 the WP did extend their guidance beyond these scenarios, when they advised Google that to overcome issues with its one-for-all privacy policy, it should provide *'an exhaustive list of the* ***types of personal data*** *processed*'[33]. This confirmed that, beyond apps and smart devices, or electronic health records, in the opinion of the WP, individuals should be informed exhaustively of the types of personal data being processed about them.

Whilst this seems promising in providing some clarity, in fact, for a number of reasons, the guidance of the WP does not provide clarity for data controllers on when and whether they need to be

---

[28] Article 29 Data Protection Working Party 'Recommendation 1/99 on Invisible and Automatic Processing of Personal Data on the Internet Performed by Software and Hardware' <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/1999/wp17_en.pdf> accessed 29 May 2016

[29] Article 29 Data Protection Working Party 'Working Document - Privacy on the Internet - An integrated EU Approach to On-line Data Protection' <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2000/wp37_en.pdf> accessed 29 May 2016

[30] Article 29 Data Protection Working Party 'Working Document on the processing of personal data relating to health in electronic health records (EHR)' <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2007/wp131_en.pdf> accessed 29 May 2016

[31] Article 29 Data Protection Working Party 'Opinion 02/2013 on apps and smart devices' <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp202_en.pdf> accessed 29 May 2016

[32] European Commission, 'First report on the implementation of the Data Protection Directive (95/46/EC)' <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52003DC0265&from=EN> accessed 9 November 2015

[33] Article 29 Data Protection Working Party 'Appendix: List of Possible Compliance Measures <http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2014/20140923_letter_on_google_privacy_policy_appendix.pdf> accessed 30 April 2016

informing individuals of the categories of personal data processed.

First, because the WP has discussed this requirement of data controllers in relation to specific scenarios (such as apps and smart devices), it is unclear whether it is only in relation to these facts that it applies. Interestingly, in relation to apps and smart devices, the WP reasons that the obligation is required because:

*'Being told **what data are being processed** is particularly important given the broad access apps generally have to sensors and data structures on the device, where such access in many cases is not intuitively obvious'*[34].

However, this justification, and the problem of unobvious and broad access, is also true of other online contexts, especially due to the increase in observed, derived, and inferred data[35]. Therefore, it would seem logical that to create transparency, this obligation should be extended to all online processing, and at least, to any other scenarios where this reasoning applies. However, until clarification is provided on this matter, the extent of this obligation remains unclear.

Second, like the hard law of the DPD, instead of constantly referring to this as a requirement to inform an individual of the 'categories' of personal data processed, the WP has used inconsistent terms in its guidance. It has referred to this requirement simultaneously, as a requirement of data controllers to inform data subjects of the 'nature of the data', the 'types of data', and of the 'categories of personal data'. Using such differing terminology to refer to this requirement, without clarifying what these terms mean (and whether they are equivalent), makes it unclear exactly what obligation the WP thinks data controllers are under.

Thirdly, although their guidance is authoritative, and highly influential, the WP holds only an advisory status, and therefore its opinions and recommendations (including these) are not legally binding. This means that if a data controller did not inform individuals of the categories of personal data it processed, it would still be for the court or the regulator to confirm that they were not fulfilling their data protection obligations. Until that confirmation, and clarification of the circumstances in which it applies, it is still not clear what the outcome will be and what obligation data controllers are under.

3.3 The United Kingdom Data Protection Act

Of course, the nature of the DPD (as a Directive) meant that it had to be implemented into each EU Member State's (MS) national law. Thus, examining these implementations could provide some clarity here. However, taking the example of the United Kingdom (UK), it is just as unclear when a data controller is under an obligation to inform the data subject of the categories of personal data they are processing.

The UK implemented the DPD through the Data Protection Act 1998[36] (DPA) and the Article 10 and

---

[34] Article 29 Data Protection Working Party 'Opinion 02/2013 on apps and smart devices' <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp202_en.pdf> accessed 29 May 2016

[35] OECD Working Party On Security And Privacy In The Digital Economy 'Protecting Privacy in a Data-driven Economy: Taking Stock of Current Thinking' <http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg%282014%293&doclanguage=en> accessed 29 May 2016

[36] Data Protection Act 1998 (UK)

11 DPD information requirements were transposed (almost verbatim) into Schedule 1, Part II 2(3) DPA. Interestingly, informing data subjects of 'the categories of data concerned' is not stated anywhere in the DPA's informational requirements, even where data is not obtained from the data subject.

Unlike the DPD, the informational requirements are only referred to once, in Schedule 1, Part II 2(3) DPA. This removed the differing 'any further information which is necessary' examples of Article 10(c) and 11(c) DPD. This is an important difference, as it was these examples that suggested that 'any further information necessary' might differ depending on whether data is obtained from the data subject or elsewhere under the DPD. Even more importantly, it was these examples that indicated that informing data subjects of the 'categories of data concerned' might even be a requirement at all. Under the DPA, the only difference between obtaining data from the individual and 'any other case' appears to be between the time of disclosure of the information, under Schedule 1 Part II, 2(1) and 2(2) DPA. Thus, the DPA provides even less clarity on when a data controller might be under an obligation to inform the data subject of the 'categories of data concerned' than the DPD.

3.4 The General Data Protection Regulation

Whilst confusing, in some ways it can be seen as quite logical that the WP might be inferring something not stated explicitly within the DPD. Indeed, it has been asserted that the OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data (which share many similarities with the DPD[37]) were '*developed primarily with 'provided data' in mind'*[38]. It is therefore not surprising that the DPD also reflects the presumption that data is collected from individuals with some degree of involvement or awareness. Following this presumption, the logical consequence is that there is no need to inform individuals of exactly what personal data is being collected, as individuals **had** to be involved or aware of data collection. The drafters could not foresee the explosion in personal digital technology that would follow the creation of the DPD, and dramatically change the personal data collection and generation practices of data controllers', rebutting this presumption. Thus, the WP may have had no choice but to try to bridge the gap between the focus of the DPD on 'provided data' and the reality of data collection as it has become, where this is just the tip of the iceberg.

Given this, and the fact that the GDPR has been heralded as the modernisation of the legal framework for data protection law within the EU, if the WP's guidance on this matter were authoritative, the logical conclusion would be that 'categories of data concerned' would be listed as a mandatory information requirement under the equivalents of **both** Article 10 and Article 11 DPD in the GDPR (Articles 13 and 14 GDPR respectively). At the very least, one would expect to see 'categories' of personal data as an example of something that an individual may need to be informed of for the processing to be fair under both Articles.

However, Article 13 GDPR (replacing Article 10 DPD) still does not mention informing data subjects of the categories of personal data at any point. The GDPR introduces new mandatory

[37] The OECD Guidelines share many similarities with the DPD, because of the close co-operation between the parties who prepared the OECD Guidelines and the Council of Europe Convention 108 (upon which the DPD is based, together with aspects of various Member State laws at the time).

[38] OECD Working Party on Security and Privacy in the Digital Economy, 'Summary of the OECD Privacy Expert Roundtable "Protecting Privacy in a Data-driven Economy: Taking Stock of Current Thinking" <http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg(2014)3&doclanguage=en> accessed 9 November 2015

information requirements under Article 13(1) GDPR, and new examples of what the data controller *might* need to inform data subjects of for processing to be fair and transparent under Article 13(2) GDPR. Yet, despite this, the GDPR still does list the categories of personal data as something data subjects may need to be informed of when personal data is obtained from them. This seems completely at odds with the WP's guidance, which has repeatedly referred to Article 10 DPD when inferring this requirement.

Interestingly, under Article 14 GDPR (which replaces Article 11 DPD), informing individuals of the categories of personal data is no longer merely an example of further information that *'might'* be necessary to ensure fair processing (Article 11(c) DPD). Under the GDPR, it is now a **mandatory informational requirement** to be given to the data subject in **every** case of data collection that is not from the data subject (Article 14(1)(d) GDPR).

Thus, on the one hand, the GDPR has increased the importance of data controllers informing data subjects of the categories of personal data. On the other hand, it is still not clear if it is **ever** an obligation for data controllers to inform individuals of this if they obtain the personal data 'from the data subject', let alone something that is mandatory in every case. Furthermore, the GDPR does not contribute any guidance on where the distinction between obtaining personal data from the data subject and from elsewhere lies.

It could be argued that the reason that the categories of personal data has not been listed in the context of personal data obtained from the data subject, is due to the fact that for some of these scenarios of data collection, individuals may be fully aware of the data they provide. However, Articles 13(4) and 14(5)(a) GDPR allow for this, stating that information is not required to be given if the data subject already has it. Therefore, it would seem more logical to make it a mandatory requirement, and then allow data controllers to rely on Article 13(4) and 14(5)(a) GDPR when necessary, rather than not include it at all. Doing so would reverse the current presumption, that individuals are aware of the data being processed about them, which is more fitting with the guidance of the WP. Furthermore, even where a data subject is aware of the categories of personal data being processed, listing them allows a data controller to attach other information to them to increase the transparency of processing, as discussed in Section 2.2.

Interestingly, taking this approach **was discussed** during the legislative process of the GDPR. The European Parliament Committee on Civil Liberties, Justice and Home Affairs (LIBE) rapporteur's draft report on amendments to the Commission's proposed GDPR[39], suggested in Amendment 126 to insert *'(aa) category of data processed'* into the (then) Article 14(1) GDPR (now Article 13(1) GDPR). The reasoning was that the GDPR:

'…can be simplified by merging information and documentation, essentially being two sides of the same coin. This will reduce administrative burdens for data controllers and make it easier for individuals to understand and exercise their rights'.

[39] European Parliament Committee On Civil Liberties, Justice And Home Affairs 'Draft Report on the proposal for a regulation of the European Parliament and of the Council on the protection of individual with regard to the processing of personal data and on the free movement of such data' <http://www.europarl.europa.eu/meetdocs/2009_2014/documents/libe/pr/922/922387/9 22387en.pdf> accessed 23 April 2016

However, the suggestion did not make its way into the LIBE Committee's Final Report[40], nor further than this in the legislative process of the GDPR.

3.5 The UK Information Commissioner's Office Guidance

In addition to examining implementations of the DPD, the WP guidance and the GDPR in search of clarity, it is also worth examining the guidance from MS supervisory authorities. However, guidance from the UK's Information Commissioner's Office (ICO) only creates further confusion.

ICO's 'Privacy notices code of practice'[41] ("the Code") is currently the leading authority on complying with the obligation to inform in the UK. A new version of the Code has recently been published, to reflect the state of the art and the impact of the GDPR. The Code aims to provide recommendations to support data controllers in drafting legally compliant, clear, and informative privacy notices.

The previous version of the Code (dated December 2010) did not mention informing data subjects of the categories of personal data processed (or 'types', or 'nature' of the data), even as an example of something that might be required in particular circumstances. Although it stated that when deciding whether to give 'any further information necessary', in the interests of fairness one must take into account the nature of the data[42], it did not state that this must be disclosed to the data subject (merely that it must be taken into account).

Yet, the lack of mention of any such requirement in the Code did not prevent ICO from advising Google that they have an obligation to inform individuals of the personal data they are processing. Following investigations into its 'one for all' privacy policy, in the undertaking with Google[43], ICO instructed them to provide:

'…clear, unambiguous and comprehensive information regarding data processing, including an exhaustive list of the ***types of data*** *processed by Google and the purposes for which data is processed'.*

This inconsistency made it even less clear when data controllers are under an obligation to provide data subjects with this information, as their guidance in practice conflicts with their guidance in the Code. Again, this could have been ICO bridging the gap between the hard law and the reality of processing today. Yet, unfortunately, the new version of the Code does not clarify the situation.

As with the previous version, there is still a strong focus on the data controller considering what information is collected internally. The Code states that to cover all the elements of fairness, an

---

[40] European Parliament Committee on Civil Liberties, Justice and Home Affairs, 'Report on the proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data' <https://polcms.secure.europarl.europa.eu/cmsdata/upload/2705202e-f65e-4a86-9de7-f2ee6d6c8d88/att_20140306ATT80606-4492192886392847893.pdf> accessed 9 May 2016

[41] Information Commissioner's Office, 'Privacy notices code of practice' <https://ico.org.uk/media/for-organisations/documents/1610/privacy_notices_cop.pdf> accessed 9 May 2016

[42] Information Commissioner's Office, 'Privacy notices code of practice' <https://ico.org.uk/media/for-organisations/documents/1610/privacy_notices_cop.pdf> accessed 9 May 2016

[43] Information Commissioner's Office, 'Google Inc. privacy policy undertaking' <https://ico.org.uk/media/action-weve-taken/undertakings/1043170/google-inc-privacy-policy-undertaking.pdf> accessed 9 November 2015

organisation will need to consider *'what information is being collected?'*[44]. It also recommends that to help decide what to include in their privacy notices, data controllers should map out how information flows through their organisation and is processed, including *'what information you hold that constitutes personal data'*[45].

Unlike the previous version, the new Code does now make it clear that a data subject will need to be informed of the categories of personal data processed. However, this only appears once, at the end of the Code, and only in relation to *'data not obtained directly from the data subject'*, reflecting the new mandatory information requirement under the GDPR (discussed in Section 3.4). There is still no discussion of when (if ever) the data subject should be informed of the categories of personal data processed if the data is obtained directly from them.

Although not using the term 'category', the new Code does use the different terminology of 'types' of data and 'the information you collect'. For example, the Code states that *'depending on the circumstances, you may decide it is beneficial to go beyond the basic requirements of the law'* and tell people the *'the links between different types of data you collect and the purposes that you use each type of data for'*[46]. Furthermore, in the Code's example of a privacy notice on a mobile screen, one of the sections to click on is called *'what information do we collect from you'?*[47].

Therefore, it is clear that the Code envisages situations where an individual must be informed of these; however, it is unclear exactly what these situations are. Although **linking** purposes to types is described as going beyond the requirements of the law, it is unclear whether 'types' should always be listed, and going beyond the law would be to link them to a purpose. Indeed, despite not listing the 'information you collect' or 'categories' or 'types' of personal data as a basic piece of information that a data controller should always include in a privacy notice[48], when discussing taking a layered approach to a notice[49], the Code states that:

'there will always be pieces of information that are likely to need to go in the top layer of a notice, such as who you are, ***what information you are collecting*** *and why you need it'*[50].

---

[44] Information Commissioner's Office, 'What should you include in your privacy notice?' <https://ico.org.uk/for-organisations/guide-to-data-protection/privacy-notices-transparency-and-control/what-should-you-include-in-your-privacy-notice/> accessed 7 October 2016

[45] Information Commissioner's Office, 'What should you include in your privacy notice?''<https://ico.org.uk/for-organisations/guide-to-data-protection/privacy-notices-transparency-and-control/what-should-you-include-in-your-privacy-notice/> accessed 7 October 2016

[46] Information Commissioner's Office, 'What should you include in your privacy notice?''<https://ico.org.uk/for-organisations/guide-to-data-protection/privacy-notices-transparency-and-control/what-should-you-include-in-your-privacy-notice/> accessed 7 October 2016

[47] Information Commissioner's Office, 'Where should you deliver privacy information to individuals?' <https://ico.org.uk/for-organisations/guide-to-data-protection/privacy-notices-transparency-and-control/where-should-you-deliver-privacy-information-to-individuals/> accessed 7 October 2016

[48] Information Commissioner's Office, 'Privacy notices under the EU General Data Protection Regulation' <https://ico.org.uk/for-organisations/guide-to-data-protection/privacy-notices-transparency-and-control/what-should-you-include-in-your-privacy-notice/> accessed 7 October 2016

[49] A 'layered approach' allows a data controller to provide the key privacy information immediately and have more detailed information elsewhere for those that want it.

[50] Information Commissioner's Office, 'Where should you deliver privacy information to individuals?' <https://ico.org.uk/for-organisations/guide-to-data-protection/privacy-notices-transparency-and-control/where-should-you-deliver-privacy-information-to-individuals/> accessed 7 October 2016

This makes it unclear exactly when a data controller is under an obligation to provide such information. It is also still unclear whether these terms all equate to the same information requirement, or whether the differing terminology reflects different information requirements.

Interestingly, on the matter of which Article applies to the processing, in relation to the GDPR, the latest version of the ICO Code states that *'there are also some differences in what you are required to provide, depending on whether you are collecting the information directly from data subjects or from a third party'*[51]. This would suggest that 'data not obtained from the data subject' means data 'collected from a third party'. This could be interpreted as meaning that data collected from a first party cookie provided by the controller would be classed as 'data obtained from the data subject' under the GDPR. If so, this would increase the importance of informing the data subject of the categories of personal data processed under Article 13 GDPR, as it is certainly not intuitively obvious to an individual what personal data is obtained from cookie. However, as it is only ICO that have elaborated in this way it remains to be seen whether this reflects general consensus under the framework. Furthermore, it is still not completely clear what 'obtained from a third party' entails and clarification with example situations under the different Articles would still prove useful here. Interestingly, the introduction of a new right under the GDPR provides potential for discussion and clarification on this matter. A new right to data portability[52] for data subjects is introduced under Article 20 GDPR. Article 20(1) GDPR provides various qualifications for the right, one of which is that the personal data must be 'provided'. Therefore, the detailed guidance expected on this right (including from the Article 29 Working Party[53]) may include discussion that clarifies the difference between data obtained from the data subject and from elsewhere or confirms this guidance from ICO.

Thus, although the new version of the Code has made some improvement on the previous version by acknowledging this as an information requirement, it is still unclear when exactly a data controller should inform the individual of the 'categories' of personal data processed and what this consists of.

3.6 Summary

As highlighted, the current and future execution of the obligation to inform makes it difficult for data controllers to understand what obligation they are under in relation to informing individuals of the categories of personal data they process.

Although the GDPR makes it clear that, a data subject must be informed of the categories when data is not obtained from them, it is still unclear exactly what this means. In addition, as the GDPR will not apply until May 2018, it is difficult for data controllers to understand what obligation **they are currently under** in relation to the DPD. Although the WP appears to confirm the position of data controllers, this may only apply in certain circumstances and the differing terminology they use makes it unclear exactly what must be done in practice to comply. Moreover, as the GDPR does not reflect the WP's guidance, it casts doubt on its applicability anyway, as neither the European

---

[51] Information Commissioner's Office, 'Privacy notices under the EU General Data Protection Regulation' <https://ico.org.uk/for-organisations/guide-to-data-protection/privacy-notices-transparency-and-control/privacy-notices-under-the-eu-general-data-protection-regulation/> accessed 7 October 2016

[52] This is a right of data subjects to receive the personal data concerning him or her in a structured, commonly used and machine-readable format, and have the right to transmit those data to another controller.

[53] Information Commissioner's Office 'The right to data portability' <https://ico.org.uk/for-organisations/data-protection-reform/overview-of-the-gdpr/individuals-rights/the-right-to-data-portability/> accessed 20 October 2016

Commission, Parliament, nor Council chose to follow the WP's guidance in its entirety.

Interestingly, in a recent Communication from the European Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, on 'Online Platforms and the Digital Single Market Opportunities and Challenges for Europe'[54], the Commission stated that:

'…large parts of the public remain apprehensive about data collection and consider that more transparency is needed. Online platforms must respond to these concerns by more effectively informing users ***what personal data is collected*** *and how it is shared and used'.*

This only confuses the position of data controllers even further, as the importance of informing users of what personal data is collected is being espoused by the Commission, but is not being clearly provided for in the law. Whilst the GDPR has contributed some definitive clarification here, it will still be unclear whether data controllers ever have to inform data subjects of the categories of personal data processed when personal data is obtained from them. It will also still be unclear just what the differences are between obtaining it from the data subject and from elsewhere are. This article urges regulators to provide direct clarification on these matters. However, such clarification will also need to confirm what a 'category' of personal data is under the obligation to inform, as Section 4 discusses.

4 How to categorise personal data?

As has been demonstrated in Section 3, it is clear that clarification of the law is required. However, understanding when a data controller is under an obligation to inform the data subject of the 'categories' of personal data is not the only issue that needs attention. Even if the law was clarified, so that:

(a) It was clear which Article data controllers processing practices were governed by; and
(b) When they are required to provide the data subjects with the *'categories of data concerned'* when collecting personal data from them (Article 13 GDPR and Article 10 DPD)

There is still the issue of exactly what a 'category of personal data' is in relation to the obligation to inform, and the question of whether any of the current approaches to categorising personal data provide meaningful information for the purposes of transparency.

4.1 What is a category of personal data?

Whilst it may initially seem obvious what a 'category' of personal data is for this obligation, upon further thought this is a valid question, the answer to which requires clarification. Indeed, an informal roundtable discussion hosted by the OECD[55], involving a cross-section of more than sixty-

[54] European Commission 'Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Online Platforms and the Digital Single Market Opportunities and Challenges for Europe (COM(2016)288)' <https://ec.europa.eu/digital-single-market/en/news/communication-online-platforms-and-digital-single-market-opportunities-and-challenges-europe> accessed 31 May 2016

[55] OECD Working Party on Security and Privacy in the Digital Economy, 'Summary of the OECD Privacy Expert Roundtable "Protecting Privacy in a Data-driven Economy: Taking Stock of Current

five privacy experts, recognised that categorising personal data could in fact be approached in numerous ways, both explicitly and implicitly. Many of these are listed in Table 1.

The notion that 'categorising' personal data is the correct approach to describing the personal data that is processed under the obligation to inform comes from the hard law of the DPD and the GDPR. However, as discussed in Section 3 (confusingly) the WP and ICO refer to this information obligation using various different additional terminologies, from the 'nature of the data' to the 'type of personal data', without clear and consistent examples of whether these terms are similar or different, and what they encompass.

Indeed, beyond the hard law, and the guidance of the WP, various other stakeholders have distinguished between personal data by identifying 'types', 'categories' and 'items', often in different ways, with differing levels of granularity. Table 1 shows how just how differently these sources (from academics, to legislation, to privacy experts) have categorised, typified and itemised personal data. This evidences the divergence in how requiring data controllers to inform data subjects of the categories of personal data being processed could be interpreted in practice without further guidance.

Table 1: Examples of different interpretations of categories, types and items of personal data

Source	Category of Data	Type of Data	Item of Data
Leon et al[56] study	Computer-related information, Demographic and preference info, Interactions with the website, Location information, Personally identifiable information	N/A	Length spent on each website page, operating system, age, gender, hobbies, country where visiting website from, name, credit card number.
Platform for Privacy Preferences (P3P) 1.0 Specification[57] Section 3.4	Physical Contact Information	N/A	Telephone number, address
	Online Contact Information	N/A	Email
	Unique Identifiers	N/A	N/A
	Purchase Information	N/A	Method of payment
	Financial Information	N/A	Credit or debit card info.

Thinking'<http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg(2014)3&doclanguage=en> accessed 9 November 2015

[56] Pedro Leon, Blase Ur, Yang Wang, Manya Sleeper, Rebecca Balebako, Richard Shay, Lujo Bauer, Mihai Christodorescu, and Lorrie Faith Cranor, 'What matters to users?: factors that affect users' willingness to share information with online advertisers' [2013] Proceedings of the Ninth Symposium on Usable Privacy and Security 7

[57] W3C 'The Platform for Privacy Preferences 1.0 (P3P1.0) Specification' <http://www.w3.org/TR/P3P/#Categories> accessed 9 November 2015

	Computer Info	N/A	IP no., domain name, browser type, operating system.
	Navigation and Click-Stream Data	N/A	Pages visited, how long users stay on each page
	Interactive Data	N/A	Queries to a search engine, or logs of account activity.
	Demographic and Socioeconomic Data	N/A	Gender, age, income
	Content	N/A	Text of email, bulletin board postings, or chat room communications
	State Management Mechanisms	N/A	N/A
	Political Information	N/A	Membership/affiliation with groups such as religious organizations, trade unions, professional associations, political parties, etc.
	Health Information	N/A	Sexual orientation, use or inquiry into health care services or products, and purchase of health care services or products.
	Preference Data	N/A	Favourite colour, musical tastes.
	Location Data	N/A	GPS position data
	Government-issued Identifiers	N/A	N/A
	Other	N/A	N/A
Allen & Overy's Guidance on Binding Corporate Rules[58]	Employment Data, Client Data	N/A	N/A
E-Privacy Directive[59]	Traffic Data	N/A	Routing, duration, time or volume of a communication, protocol used, location

[58] Allen & Overy, 'Binding Corporate Rules' <http://www.allenovery.com/SiteCollectionDocuments/BCRs.pdf > accessed 9 May 2016

[59] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) [2002] OJ L201/37

			of the terminal equipment of the sender or recipient, network on which the communication originates or terminates, beginning, end or duration of a connection
	Location Data	N/A	Latitude, longitude and altitude of the user's terminal equipment, direction of travel, level of accuracy of the location information, the identification of the network cell in which the terminal equipment is located at a certain point in time, time the location information was recorded
Data Protection Directive 95/46/EC[60]	Racial or ethnic origin Political opinions Religious or philosophical beliefs Trade-union membership Health or sex life data	N/A	N/A
OECD Privacy Expert Roundtable[61]	*(Categorisations in relation to the sensitivity of the data)* Health Data Ethnic Origin	N/A	N/A
	(Categorisations in relation to the subject of the data)	N/A	N/A

[60] Directive 95/46/EC of the European Parliament and of the Council of on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281/31

[61] OECD Working Party on Security and Privacy in the Digital Economy, 'Summary of the OECD Privacy Expert Roundtable "Protecting Privacy in a Data-driven Economy: Taking Stock of Current Thinking' <http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg(2014)3&doclanguage=en> accessed 9 May 2016

	Employee Data Minor's Data Non-citizens Data		
	(Categorisations in relation to the context in which the data is being processed) Electronic Communications Data, Credit Reporting Data, Archival Data, Social Security Administration Data	N/A	N/A
	(Categorisations in relation to the degree of identifiability) Identifying Data De-identified Data Anonymous Data Pseudonymous Data	N/A	N/A
	(Categorisations in relation to how the data has been collected) Directly collected data Indirectly collected data	N/A	N/A
	(Categorisations in relation to the manner in which the data originated) Provided Data Observed Data Derived Data Inferred Data	N/A	N/A
General Data Protection Regulation[62]	Data revealing: Racial or ethnic origin	N/A	N/A

[62] 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) http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.119.01.0001.01.ENG&toc=OJ:L:2016:119:TOC accessed 29 May 2016

	Political opinions, Religious or philosophical beliefs, Trade-union membership,		
	Genetic data, Biometric data Data concerning health Data concerning a natural person's sex life or sexual orientation	N/A	N/A
UK Data Protection Register	N/A	Personal Details Family, Lifestyle and Social Circumstances Financial Details Employment and Education Details Goods or Services Provided	N/A
MyDex White Paper '*The Case for Personal Information Empowerment: The rise of the personal data store*'[63]	N/A	Data that identifies	Name, address
	N/A	Data conferred by other parties	Passport number, my credit reference rating
	N/A	Information gathered by me	Search and research results
	N/A	Data generated by my dealings with other parties	Transaction and interaction records)
	N/A	Information created by me	My plans, my preferences

Thus, with so many different approaches, even if the law were to be clarified so that a data controller could understand **when** they are under an obligation to inform the data subject of the categories of personal data they process, without further clarification, it is still unclear exactly what information they would need to be providing. Therefore, clarification is also required on which of these approaches is the one referred to in relation to the obligation to inform.

4.2 Are any of these categorisations useful?

In clarifying what a 'category of personal data' is for the purposes of the obligation to inform, it is important to consider whether any of these categorisations actually provide meaningful information that will increase the transparency of data processing for data subjects, allowing them to assess the compliance and trustworthiness of the data controller. After all, this is the purpose of this obligation,

[63] Mydex,'The Case for Personal Information Empowerment: The rise of the personal data store' < <https://mydex.org/wp-content/uploads/2010/09/The-Case-for-Personal-Information-Empowerment-The-rise-of-the-personal-data-store-A-Mydex-White-paper-September-2010-Final-web.pdf> accessed 9 May 2016

and it is achieved by redressing the balance of information between the data subject and the data controller. Whilst each of the categorisations in Table 1 have their use, they are not necessarily useful for making data processing practices more transparent on their own. The next sections discuss some of the different approaches to categorising personal data. In doing so, it explains why each of these on their own are insufficient to increase the transparency of personal data processing to a level that equates the information available to subjects to that of data controllers, allowing them to assess compliance and trustworthiness of the controller.

4.2.1 Categorising personal data in relation to identifiability

Personal data can be categorised in relation to the degree of identifiability e.g. by distinguishing between identifying data, de-identified data, anonymous data and pseudonymous data. Informing data subjects of which of these categories are processed may be useful for helping them ascertain when data protection laws apply. However, to check whether the data controller is compliant, including whether security obligations are complied with, more information will be required. This approach alone does not help individuals understand exactly **what** is being collected or how it will be processed, to allow them to make subjective and granular decisions about these aspects. Without this granularity, although data subjects may know when data protection laws apply, they will not be able to assess compliance.

4.2.2 Categorising in relation to sensitivity

Under the DPD and GDPR, the only categories of personal data that are clearly defined (as Table 1 shows) are the 'special categories'. Both the DPD and the GDPR provide definitions of 'categories' of personal data they deem 'special' and thus warranting further protection under the framework (Article 8(1) DPD and Article 9(1) GDPR). Thus, when the example or requirement of being informed of the 'categories of data' is referred to under the obligation to inform, it could be interpreted that the requirement refers to informing the individual of whether 'special categories' of personal data are processed.

Indeed, distinguishing between 'personal data' and 'sensitive personal data' and informing individuals of which category a data controller processes could prove useful for data subjects. It could help them understand the sensitivity of the data in question and help them to keep a check on data controller compliance with other obligations under the framework in relation to processing special categories. This approach could also help a data subject to assess the risk of the personal data processing. However, despite these benefits, to make data processing transparent, categories need to have a lower level of abstraction than just 'personal data' and 'sensitive personal data'.

Therefore, the requirement could be interpreted as listing the specific categories of personal data being processed. However, although the DPD and GDPR provide a lower level of abstraction for categories deemed 'special' or 'sensitive' i.e. data revealing racial or ethnic origin or political opinions etc., they do not provide the equivalent for 'non-sensitive' personal data.

Furthermore, being informed of these categories does not make it clear exactly what personal data is collected, but simply that how it is processed places it in a category of 'sensitive' or 'special' data. For example, informing data subjects that a data controller processes 'data revealing ethnic origin' does not make it clear to the individual whether it is their provided ethnic origin is being processed for this purpose or whether, assumptions are being made on their name or residential status. Therefore, this approach to categorisation alone would not support an individual in understanding what is being collected and processed exactly, yet it would give them a general sense of the personal

data is being derived or inferred (the importance of context in relation to personal data processing is discussed further in Section 4.2.3).

4.2.3 Other categorisations, data types and taxonomies

In providing a lower level of abstraction for 'non-sensitive' personal data, some of the other categorisations in Table 1, such as those of Leon et al[64], P3P[65], and the e-Privacy Directive[66], could prove useful. Focusing on a data controller informing the individual of these could see a move towards the creation of a taxonomy of personal data, as called for by the World Economic Forum[67]. Yet, this approach would face various issues, making it inappropriate as an approach to categorising personal data under the obligation to inform.

First, there is the issue of creating a taxonomy that is simultaneously able to:

- Accommodate new forms of personal data as new technologies emerge;
- Remain simple enough for data controllers and data subjects to comprehend; and
- Be able to deal with personal data that belongs to more than one category.

Indeed, one of the criticisms of the Platform for Privacy Preferences Project (P3P)[68] (and its lack of adoption) was because their approach to categorising data was too complex, even for webmasters[69]. Yet, their taxonomy only included seventeen 'data types'[70]. Given the rate at which new forms of data are being created and utilised, such taxonomy is likely to become confusing quickly. This would suggest that any approach to categorisation would need to include far fewer possible categories than this.

Second, is the issue of deciding on the granularity (or level of abstraction) of the categories, which requires making trade-offs between specificity and practicality. An example of a P3P category is 'computer information', yet even this is a wide category, which does not make it intuitively obvious what it includes. This means that exactly what is being collected is still not transparent. Whilst increasing the granularity, to state that the data subject's 'IP Address' is processed may increase transparency, it may also result in even longer privacy notices and cognitive overload, given the

---

[64] Pedro Leon, Blase Ur, Yang Wang, Manya Sleeper, Rebecca Balebako, Richard Shay, Lujo Bauer, Mihai Christodorescu, and Lorrie Faith Cranor, 'What matters to users?: factors that affect users' willingness to share information with online advertisers' [2013] Proceedings of the Ninth Symposium on Usable Privacy and Security 7

[65] W3C 'The Platform for Privacy Preferences 1.0 (P3P1.0) Specification' <http://www.w3.org/TR/P3P/#Categories> accessed 9 November 2015

[66] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) [2002] OJ L201/37

[67] World Economic Forum, 'Rethinking Personal Data: A New Lens for Strengthening Trust' <http://www3.weforum.org/docs/WEF_RethinkingPersonalData_ANewLens_Report_2014.pdf> accessed 9 April 2016

[68] A protocol allowing websites to declare their intended use of information they collect about web browser users.

[69] Ari Schwartz 'Looking Back at P3P: Lessons for the Future' <http://ec.europa.eu/justice/news/events/workshop_pets_2009/presentations/SCHWARTZ_Ari_paper.pdf> accessed 31 May 2015

[70] Ari Schwartz 'Looking Back at P3P: Lessons for the Future' <http://ec.europa.eu/justice/news/events/workshop_pets_2009/presentations/SCHWARTZ_Ari_paper.pdf> accessed 31 May 2015

amount of different data types there already are in existence, let alone those to be created (especially as even webmasters struggled with just seventeen).

Third, is the issue that these categorisations focus only on the collection aspect of the data processing i.e. what personal data is collected? This is only the first step of making the data processing transparent, as ignores the various affordances of personal data[71] and the role that context plays in personal data processing. This fails to make what might, and what is, going to be done with the personal data transparent.

For example, in 2013[72] a study showed that Facebook 'likes' could be processed to predict a wide range of other personal data (much of which would be deemed 'sensitive personal data') such as sexuality and political views. Here, a focus on informing the individual about the collected data alone may have resulted in simply telling them that their 'Facebook likes' are collected, or (in taking a 'category approach') simply telling them that their 'interactions on the website' are collected. However, neither of these would have made the affordances of this data clear to the individual. It does not help them understand everything that could, or will, be derived or inferred from this personal data. It also ignores the fact that the same data may or may not be personal data depending on the context[73]. Describing what might, and what is going to be done with personal data is a basic element of data protection regulation. This is embodied in the principle of purpose limitation, as discussed in Section 2.1.

Thus, using categorisations that focus only on the collection aspect of the data processing (and not what might, and what is going to be done with it) ignores a basic element of data protection regulation, and reduces the transparency of data processing significantly. Interestingly, combining this approach and the current approach to special categories discussed in Section 4.2.2 could prove useful here in making what is collected and what it is being processed to reveal transparent. Any approach to categorisation should acknowledge the vital importance of context and consider the lifecycle of the personal data and the purposes that will be applied to it in order to categorise it also, rather than simply its status at the time of collection alone.

4.2.4 Categorising in relation to the manner in which the data originated

In relation to the last issue of focusing beyond the point of collection of the personal data alone, the categories produced by the OECD's Privacy Expert Roundtable, of 'provided', 'observed', 'derived' and 'inferred' personal data[74] could also prove useful. They could be used to make the individual aware of whether the personal data collected will remain in that form, or whether it will be used to create or predict other personal data. However, again these would need to be used in combination with another approach, as there is still the issue of having a lower level of abstraction, to help

---

[71] 'Affordance' here means that a specific 'data type' can be processed to derive and infer further 'data types'.
[72] Michal Kosinski, David Stillwell, Thore Graepel.'Private traits and attributes are predictable from digital records of human behavior' [2013] Proceedings of the National Academy of Sciences, 110(15), pp.5802-5805
[73] Article 29 Data Protection Working Party 'Opinion 4/2007 on the concept of personal data' <http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf> accessed 31 May 2016
[74] World Economic Forum, 'Rethinking Personal Data: A New Lens for Strengthening Trust' <http://www3.weforum.org/docs/WEF_RethinkingPersonalData_ANewLens_Report_2014.pdf> accessed 9 April 2016

individuals understand exactly **what** is being 'observed', 'derived' and 'inferred'. This would support them in making subjective choices about this and/or acting as a check on data controllers.

4.3 IP addresses

The example of IP Addresses[75] highlights the need for a better approach to categorising personal data under the obligation to inform, one that encompasses the benefits of the individual approaches to categorisation. An IP address can be used for many different purposes and could be classed as both 'provided' and 'observed' data (depending on how it is collected). However, an IP address can also be used to 'derive' an individual's location. Based on this location, predictions can then be made, and more personal data 'inferred' (possibly based on the personal data of other individuals who share that location). However, a data controller could just be collecting a data subject's IP address and doing nothing further with it. Currently, using any of the approaches to categorisation analysed in this article alone would not make it clear simultaneously what is being collected, whether it will be processed further and the limits on how it will be processed.

4.4 How does Google handle this?

Of course, if despite the lack of clarity in the legal framework, data controllers were describing 'categories of personal data' in a way that makes data processing transparent, then these issues would not be as pressing. However, if we look at Google's privacy policy[76] alone (dated 29 August 2016), as a data controller who has received instructions from both the WP and ICO on this matter (as discussed in Sections 3.2 and 3.5), and as the number one ranked website in the world[77], they do not categorise personal data in a way that makes processing transparent and allows a data subject to assess Google's compliance and trustworthiness either.

Their policy has a section called 'Information that we collect' and within this there are two sub-headings, 'Information you give us' and 'Information we get from your use of our services'. The much larger amount of text under the latter heading alone, confirms that 'Information you give us' is merely the tip of the iceberg in relation to data collection.

Under the heading 'Information we get from your use of our services', there are six further subsections: Device information, Log information, Location information, Unique application numbers, Local storage and Cookies and similar technologies. Within each of these categories, more specific information is provided on what is collected. However, this still does not make data processing transparent for a number of reasons.

First, it is not an exhaustive list of the types of information provided, as the words 'this includes' and 'for example' (among others to the same effect) within this list confirm.

Second, these categories focus more on the source of the information (e.g. cookies) than what is actually collected, with only non-exhaustive examples of this provided. Whilst understanding the source of personal data can support an individual in having control over the flow of personal data (e.g. by stopping using an app), this alone does not increase the transparency of personal data processing.

---

[75] An IP address is a unique string of numbers separated by full stops that identifies each computer using the Internet Protocol to communicate over a network.

[76] Google 'Google Privacy and Terms' <https://www.google.co.uk/intl/en/policies/privacy/?fg=1> accessed 23 September 2016

[77] Alexa, 'The top 500 sites on The Web' <http://www.alexa.com/topsites> accessed 1 June 2016

Finally, the focus is only on what is collected, or in other words, what is 'provided' and 'observed'. As discussed in Section 4.2.3, this focus ignores a basic element of data protection law and the important role context and processing play. There is no mention of whether this information is used to derive or infer any further information, and because it is non-exhaustive and not linked to a purpose, this is difficult for the data subject to ascertain.

Therefore, in practice data subjects are not being provided with the information they need for their data processing to be transparent by data controllers. This makes the need for clarification on this even more pressing. It is difficult to reprimand Google and other data controllers for lacking an appropriate approach to categorising personal data in their privacy policies when the law is in desperate need of clarification.

5 Conclusion

This article has highlighted that despite the benefits of categorising personal data, a coherent and consistent approach to doing so under the obligation to inform has not emerged. It has demonstrated the uncertainty over both what a 'category' of personal data is in relation to this obligation and when this information must be provided. Ultimately, this uncertainty results in reduced transparency for data subjects and confusion for data controllers regarding their legal obligations, defeating the purpose of the obligation to inform. This article highlights these issues and calls for clarification on them. This article also posits that a new approach to categorising personal data is required, given the deficiencies of the current approaches in increasing transparency on their own.

A supervisory authority could take advantage of this lack of clarity in the correct approach to categorisation under the obligation to inform and clarify a new approach within a code of conduct. Indeed, the recent version of the ICO 'Privacy Notices Code of Practice' states that ICO will consider producing further guidance on the obligation's individual information requirements under the GDPR[78]. This could include further information on *'categories of personal data'*, taking advantage of the contributions of this article. Whilst proposing and describing a new approach to categorisation is beyond the scope of this article, it is important to understand how the benefits of categorisation can best be achieved. Therefore, this article provides a discussion of the benefits to be achieved, and issues to be avoided, in any proposal of a new approach.

The categorisation of personal data has clear benefits for making personal data processing more transparent for data subjects. With the increased importance of transparency under the GDPR, it is important that these issues with the current approach to categorisation under the obligation to inform are overcome, for the benefit of data subjects, controllers, and for the efficacy of the EU data protection framework as a whole.

Acknowledgements

This research was funded by the Research Councils UK Digital Economy Programme, Web Science Doctoral Training Centre, University of Southampton. EP/G036926/1

Bibliography

[78] Information Commissioner's Office, 'Privacy notices under the EU General Data Protection Regulation' <https://ico.org.uk/for-organisations/guide-to-data-protection/privacy-notices-transparency-and-control/privacy-notices-under-the-eu-general-data-protection-regulation/> accessed 7 October 2016

Aleecia M McDonald and Lorrie Faith Cranor, 'The cost of reading privacy policies' [2008] *ISJLP* 4: 543

Alexa, 'The top 500 sites on The Web' <http://www.alexa.com/topsites> accessed 1 June 2016

Allen & Overy, 'Binding Corporate Rules' <http://www.allenovery.com/SiteCollectionDocuments/BCRs.pdf > accessed 9 May 2016

Ari Schwartz 'Looking Back at P3P: Lessons for the Future' <http://ec.europa.eu/justice/news/events/workshop_pets_2009/presentations/SCHWARTZ_Ari_paper.pdf> accessed 31 May 2015

Article 29 Data Protection Working Party 'Appendix: List of Possible Compliance Measures <http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2014/20140923_letter_on_google_privacy_policy_appendix.pdf> accessed 30 April 2016

Article 29 Data Protection Working Party 'Opinion 02/2013 on apps and smart devices' <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp202_en.pdf> accessed 29 May 2016

Article 29 Data Protection Working Party 'Opinion 03/2013 on purpose limitation' <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp203_en.pdf> accessed 20 October 2016

Article 29 Data Protection Working Party 'Opinion 4/2007 on the concept of personal data' <http://ec.europa.eu/justice/policies/privacy/docs/wpdocs/2007/wp136_en.pdf> accessed 31 May 2016

Article 29 Data Protection Working Party 'Recommendation 1/99 on Invisible and Automatic Processing of Personal Data on the Internet Performed by Software and Hardware' <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/1999/wp17_en.pdf accessed 29 May 2016>

Article 29 Data Protection Working Party 'Working Document - Privacy on the Internet - An integrated EU Approach to On-line Data Protection' <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2000/wp37_en.pdf> accessed 29 May 2016

Article 29 Data Protection Working Party 'Working Document on the processing of personal data relating to health in electronic health records (EHR)' <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2007/wp131_en.pdf > accessed 29 May 2016

BEUC 'A Comprehensive Approach on Personal Data Protection in the European Union, European Commission's Communication: The European Consumers' Organisation's response' <http://ec.europa.eu/justice/news/consulting_public/0006/contributions/organisations/b euc_en.pdf> accessed 20 April 2016

Data Protection Act 1998 (UK)

Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) [2002] OJ L201/37

Directive 95/46/EC of the European Parliament and of the Council of on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281/31

Eleanor Rosch and Barbara Lloyd, (eds) *Cognition and categorization* 27-48. (1978, Lawrence Erlbaum)

European Commission 'Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Online Platforms and the Digital Single Market Opportunities and Challenges for Europe (COM(2016)288)' <https://ec.europa.eu/digital-single-market/en/news/communication-online-platforms-and-digital-single-market-opportunities-and-challenges-europe> accessed 31 May 2016

European Commission, 'First report on the implementation of the Data Protection Directive (95/46/EC)' <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52003DC0265&from=EN> accessed 9 November 2015

European Parliament Committee on Civil Liberties, Justice and Home Affairs, 'Report on the proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data' <https://polcms.secure.europarl.europa.eu/cmsdata/upload/2705202e-f65e-4a86-9de7-f2ee6d6c8d88/att_20140306ATT80606-4492192886392847893.pdf> accessed 9 May 2016

European Parliament Committee On Civil Liberties, Justice And Home Affairs 'Draft Report on the proposal for a regulation of the European Parliament and of the Council on the protection of individual with regard to the processing of personal data and on the free movement of such data' <http://www.europarl.europa.eu/meetdocs/2009_2014/documents/libe/pr/922/922387/922387en.pdf> accessed 23 April 2016

European Union Agency for Fundamental Rights, Council of Europe and the Registry of the European Court of Human Rights 'Handbook on European data protection law' <http://www.echr.coe.int/Documents/Handbook_data_protection_ENG.pdf> accessed 29 May 2016

Gerrit-Jan Zwenne, Diluted Privacy Law (April 12, 2013). Available at SSRN: https://ssrn.com/abstract=2488486

Google 'Google Privacy and Terms' <https://www.google.co.uk/intl/en/policies/privacy/?fg=1> accessed 23 September 2016

Information Commissioner's Office, 'Google Inc. privacy policy undertaking' <https://ico.org.uk/media/action-weve-taken/undertakings/1043170/google-inc-privacy-policy-undertaking.pdf> accessed 9 November 2015

Information Commissioner's Office, 'Overview of the General Data Protection Regulation (GDPR)' <https://ico.org.uk/media/for-organisations/data-protection-reform/overview-of-the-gdpr-1-0.pdf> accessed 29 September 2016

Information Commissioner's Office, 'Privacy notices code of practice' <https://ico.org.uk/media/for-organisations/documents/1610/privacy_notices_cop.pdf> accessed 9 May 2016

Information Commissioner's Office, 'Privacy notices under the EU General Data Protection Regulation' <https://ico.org.uk/for-organisations/guide-to-data-protection/privacy-notices-transparency-and-control/what-should-you-include-in-your-privacy-notice/> accessed 7 October 2016

Information Commissioner's Office, 'What should you include in your privacy notice?' <https://ico.org.uk/for-organisations/guide-to-data-protection/privacy-notices-transparency-and-control/what-should-you-include-in-your-privacy-notice/> accessed 7 October 2016

Information Commissioner's Office, 'Where should you deliver privacy information to individuals?' <https://ico.org.uk/for-organisations/guide-to-data-protection/privacy-notices-transparency-and-control/where-should-you-deliver-privacy-information-to-individuals/> accessed 7 October 2016

Information Commissioner's Office 'The right to data portability' <https://ico.org.uk/for-organisations/data-protection-reform/overview-of-the-gdpr/individuals-rights/the-right-to-data-portability/> accessed 20 October 2016

Michal Kosinski, David Stillwell, Thore Graepel.'Private traits and attributes are predictable from digital records of human behavior' [2013] Proceedings of the National Academy of Sciences, 110(15), pp.5802-5805

Mistale Taylor 'Safeguarding the Right to Data Protection in the EU, 30th and 31st October 2014, Paris, France' [2015] Utrecht J. Int'l & Eur. L., 31, 145.

Mydex,'The Case for Personal Information Empowerment: The rise of the personal data store' <https://mydex.org/wp-content/uploads/2010/09/The-Case-for-Personal-Information-Empowerment-The-rise-of-the-personal-data-store-A-Mydex-White-paper-September-2010-Final-web.pdf> accessed 9 May 2016

OECD Working Party on Security and Privacy in the Digital Economy, 'Summary of the OECD Privacy Expert Roundtable "Protecting Privacy in a Data-driven Economy: Taking Stock of Current Thinking" <http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=dsti/iccp/reg(2014)3&doclanguage=en> accessed 9 November 2015

OECD, ‘The Evolving Privacy Landscape: 30 Years After the OECD Privacy Guidelines’ <http://www.oecd-ilibrary.org/docserver/download/5kgf09z90c31.pdf?expires=1477319599&id=id&accname=guest&checksum=3CCA414864E0A35F15FA336CB015E5B8> accessed 20 October 2016

Oxford English Mini Dictionary (First published 1981, Oxford University Press, 2011) 82

Pedro Leon, Blase Ur, Yang Wang, Manya Sleeper, Rebecca Balebako, Richard Shay, Lujo Bauer, Mihai Christodorescu, and Lorrie Faith Cranor, ‘What matters to users?: factors that affect users' willingness to share information with online advertisers’ [2013] Proceedings of the Ninth Symposium on Usable Privacy and Security 7

Perri 6, *The future of privacy: Volume 1 Private life and public policy* (Demos 1998)

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) <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2016.119.01.0001.01.ENG&toc=OJ:L:2016:119:TOC> accessed 29 May 2016

W3C ‘The Platform for Privacy Preferences 1.0 (P3P1.0) Specification’ <http://www.w3.org/TR/P3P/#Categories> accessed 9 November 2015

World Economic Forum, ‘Rethinking Personal Data: A New Lens for Strengthening Trust’ <http://www3.weforum.org/docs/WEF_RethinkingPersonalData_ANewLens_Report_2014.pdf> accessed 9 April 2016