The idea of 'emergent properties' in data privacy: towards a holistic approach

Samson Y. Esayas*

ABSTRACT

'The whole is more than the sum of its parts.'¹

This article applies lessons from the concept of 'emergent properties' in systems for data privacy law. This concept, rooted in the Aristotelian dictum 'the whole is more than the sum of its parts', where the 'whole' represents the 'emergent property', allows systems engineers to look beyond the properties of individual components of a system and understand the system as a single complex. Applying this concept, the article argues that the current European Union data privacy rules focus on individual processing activity based on a specific and legitimate purpose, with little or no attention to the totality of the processing activities—ie the whole—based on separate purposes. This implies that when an entity processes personal data for multiple purposes, each processing must comply with the data privacy principles separately, in light of the specific purpose and the relevant legal basis.

This (atomized) approach is premised on two underlying assumptions: (i) distinguishing among different processing activities and relating every piece of personal data to a particular processing is possible, and (ii) if each processing is compliant, the data privacy rights of individuals are not endangered. However, these assumptions are untenable in an era where companies process personal data for a panoply of purposes, where almost all processing generates personal data and where data are combined across several processing activities. These practices blur the lines between different processing activities and complicate attributing every piece of data to a particular processing. Moreover, when entities engage in these practices, there are privacy interests independent of and/or in combination with the individual processing activities. Informed by the discussion about emergent property, the article calls for a holistic approach with enhanced responsibility for certain actors based on the totality of the processing activities and data aggregation practices.

* Doctoral research fellow at the Norwegian Research Center for Computers and Law (NRCCL), Department of Private Law, University of Oslo, Oslo. E-mail: s.y.esayas@jus.uio.no. This work is financed by the University of Oslo and partly supported by the SIGNAL project (Security in Internet Governance and Networks: Analysing the Law), which is jointly funded by the Norwegian Research Council and UNINETT Norid AS. The author is grateful to Lee A. Bygrave and Inger B. Ørstavik for their valuable comments on several drafts. The author would also like to thank Worku Urgessa for his comments on earlier draft of this article. However, the usual disclaimer applies.

¹ The quotation is often attributed to Aristotle, although others are also credited for it. See Daniel Solove, *Understanding Privacy* (Harvard UP 2008) 117.

© The Author (2017). Published by Oxford University Press. All rights reserved.

For Permissions, please email: journals.permissions@oup.com.

_ 139

International Journal of Law and Information Technology, 2017, 25, 139–178

doi: 10.1093/ijlit/eaw015

Advance Access Publication Date: 15 February 2017

Article

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

INDIVIDUALISTIC APPROACH TO DATA PRIVACY

I use the term 'individualistic' in the sense that in dealing with a system built from several components, one emphasizes the constituent (individual) parts of the system as opposed to the system as a collective whole. In the context of European Union (EU) data privacy rules, this issue of 'individualistic nature' arises when an entity processes personal data for multiple purposes. The use of the term 'purposes' (plural) under Article 6(1(b)) of the Directive² or Article 5(1(b)) of the newly adopted Regulation³ implies that personal data can be collected for more than one purpose. In such cases, the Article 29 Working Party requires that 'each separate purpose should be specified in enough detail to be able to assess whether collection of personal data for this purpose complies with the law' and 'the data quality requirements must be complied with separately for each purpose'.⁴

The main argument in this article is that this (individualistic) approach is inadequate in an era where companies process personal data for manifold purposes, where almost every use of the service generates personal data, and where data are

combined across several processing activities. This is partly because the individualistic approach is based on the underlying assumption that there are well-delineated, distinct processing activities serving distinct purposes, with every piece of data fitting into those delineated individual boxes of processing activities. However, in light of the increasing commercial value of personal data and big data practices, this assumption is a half-truth at best. Furthermore, the account of the individualistic approach partially signifies that if the individual processing is compliant, the data privacy rights of individuals are not endangered. However, as shown throughout the article, this might not always be the case. By drawing comparisons from the concept of 'emergent property', I demonstrate that when entities process personal data for a wide array of purposes and combine data across these processing activities, there are privacy interests independent of and/or in combination with the individual processing activities.⁵ These interests include overexposure of the individual, loss of transparency and accountability, and loss of practical obscurity. First, however, I wish to clarify what the individualistic approach is.

The compliance assessment lifespan for a certain processing activity commences by determining the specific purpose of collecting personal data.⁶ Upon determining

² Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 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 (hereinafter, the Directive).

³ 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 [2016] OJ L119/1 (hereinafter, GDPR).

⁴ Article 29 Data Protection Working Party, 'Opinion 03/2013 on Purpose Limitation' [2013] 12 and 16 (hereinafter, WP203) (emphasis added).

⁵ For more on 'Values and Interests Safeguarded by Data Protection Law', see Lee Bygrave, *Data Protection Law: Approaching Its Rationale, Logic and Limits* (Kluwer Law International 2002) ch 7; see also section 'Privacy interests in wholeness and sum'.

⁶ In fact, one could argue that this is not how it actually works in practice, because the essence of big data lies in collecting large volumes of data from all possible sources (instruments, sensors, Internet transactions, e-mail, video, click streams), analysing such data with the support of advanced analytical tools and techniques, and the subsequent reuse of such data. This implies that often the purpose of such data is determined after extensive analysis of the data in combination with other data. This idea somehow defies

140 Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

the purpose, one needs to find an appropriate legal basis from among the six alternative legal grounds under Article 7 (or Article 8(2) if it is sensitive personal data) of the Directive that suits the specified purpose.⁷ The determination of the purpose and relevant legal basis creates a set box for that particular processing activity, as shown in the middle of [Figure 1](#), which becomes the epicentre of the application of the core data privacy principles.⁸

In other words, a regulator or an internal auditor who is interested in assessing the compliance of that particular processing activity should be able to identify the specific purpose for which the personal data were collected; if there is a legitimate basis and that this legal basis suits the particular purpose; if the data collected were the minimum necessary for the purpose; if the data collected were accurate and up-to-date for that particular purpose; if the personal data are not stored for longer than necessary to achieve the purpose; and if the processing of such data is fair (ie it does not exceed the expectations of individuals). Apart from those principles anchored to Article 6 of the Directive, the assessment should consider whether appropriate technological and organizational measures are in place 'to protect personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorized disclosure or access' (data security principle)⁹; and if the processing involves a transfer of personal data to a third country, in which case the requirements under Article 26 demand compliance.¹⁰ [Figure 1](#) depicts this atomized way of assessing compliance with the rules.

When an entity processes personal data for several purposes, this process repeats itself for as many distinct purposes as there are. The distinct processing activities can either be based on different purposes and different legal grounds under Article 7 (Article 8(2)) or different purposes, but the same legal ground, which are unrelated.

¹¹ For example, an entity might be involved in processing personal data for

marketing purposes (and using consent as a legal basis); or processing a payment in a transaction (and using the 'performance of a contract entered into with the data subject' as a legal basis); or processing data for taxation purposes (and using compliance with legal obligations as a legal basis). Similarly, an entity might process personal data based on the same legal basis, but for different purposes. For example, the determination of a specific purpose in advance. See Lokke Moerel, 'Big Data Protection', How to Make the Draft EU Regulation on Data Protection Future Proof (Tilburg University 2014) 7. However, this claim is well in line with the overall argument that the current data privacy rules are inadequate for dealing with emerging data processing practices.

7 Art 7 of the Directive provides that personal data may be processed only if (i) the data subject gives unambiguous consent; (ii) processing is necessary for the execution of a contract; (iii) processing is necessary for compliance with a legal obligation; (iv) processing is necessary for protecting the vital interests of the data subject; (v) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority; and (vi) processing is necessary for the purposes of the controller's legitimate interests.

8 These include the lawfulness and fairness principle (Art 6(1(a)), purpose limitation principle (Art 6(1(b))), data minimization principle (Art 6(1(c))), accuracy or data quality principle (Art 6(1(d))) and the principle regarding storage limitation (Art 6(1(e))). The data security principle can be found under Art 17(1) of the Directive.

9 Directive, Art 17.

10 Directive, Art 25.

11 For discussions on 'related' processing operations, see WP203 (n 4) 16 and examples in Annex 3.

Emergence in data privacy: a holistic approach _ 141

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

consent can be used to process personal data, to serve advertising as well as research purposes. Each of those processing activities based on a specific purpose and a distinct legitimate ground forms a set box, as represented by the box at the centre of Figure 1 above. This means that when an entity is processing personal data for more than one purpose, each box should comply with the data privacy principles separately.

12 In the above figure, those principles are represented by the questions embedded in the boxes attached to the central box. In order to apply the principles to each box, one must identify the different set boxes based on the specific purpose and legitimate basis and the data that belong to within each box. This represents an individualistic view of the processing; in other words, the assessment of compliance focuses on the individual processing based on a specific purpose and distinct legal basis, regardless of the total number of processing activities and data aggregation practices across the different processing operations.¹³

Figure 1. Individualistic view of data privacy compliance assessment.

12 The Working Party makes it clear that 'If personal data are processed for several purposes, all requirements of Art. 6 apply to each purpose separately'; *ibid* 16.

13 The discussions about the 'individualistic approach' should not be confused with emerging discussions about the focus of data privacy rules on 'identifiable individuals', which have been criticized for

142 _ Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

Before proceeding further, I would like to mention two points of caution. First, it can be argued that in most cases, the purpose limitation principle will not allow data aggregation across different processing activities based on distinct purposes. It goes without saying that aggregation constitutes processing under the EU rules and requires a legitimate basis of its own when conducted on processing operations based on distinct purposes. This means that if an entity aggregates data across different processing activities based on separate purposes, without having a legitimate basis, it would constitute a breach of the rules. However, this task is complicated by the reliance of many entities on consent as a basis for processing data including the aggregation practices. Once the user consents to such aggregation practices, the line between the different individual processing activities and the data within each box starts to disappear, complicating the application of the data privacy rules to a specific box. The second point relates to the flexible tools that Data Protection Authorities (DPAs) have to consider the relevant interests under each circumstance. There is some evidence that the authorities are not restricted to looking at the individual processing activities but, rather, at the overall privacy interests of individuals.¹⁴ In this regard, it bears mentioning the flexibility that some principles, such as the fairness

principle, provide in terms of looking beyond individual processing activities. However, these kinds of encouraging moves towards a holistic approach need to be strengthened by relevant tools and a theoretical framework for their consistent application. The aim of this article is to contribute towards achieving such an objective. To this end, I use the concept of emergent property to shed light on how to understand a system involving several processing operations as a single complex. The motivation for using the concept of emergent property in data privacy law is two-fold. First, the growing importance of personal data for commercial purposes is fuelling the desire to collect and amass as much personal data, both through legitimate (eg acquisitions) and illegitimate (eg deliberate deception) mechanisms, leading to the emergence of companies processing personal data for a panoply of purposes. Secondly, big data practices represent substantial value in repurposing, recombining and multi-purposing of personal databases. A salient example is Google's, now under the umbrella of parent company Alphabet,¹⁵ aggressive expansion to a broad array of new product areas ranging from e-mail, search, map service, video-sharing, social network, mobile operating system, payment service, etc. An investigation from 2013 by the Spanish Data Protection Authority shows that Google collects personal data through nearly 100 'consumer-facing products or services'.¹⁶ Moreover, since their overlooking the dangers of group privacy based on profiling. See Alessandro Mantelero, 'Personal Data for Decisional Purposes in the Age of Analytics: From an Individual to a Collective Dimension of Data Protection' (2016) 32 (2) Computer Law & Security Review 238. Similarly, some authors refer to the individualistic approach in relation to the focus of the EU rules on the individual's ability to control his/her personal data. See eg Christophe Lazaro and Daniel Le Me'tayer, 'Control over Personal Data: True Remedy or Fairy Tale?' (2015) 12 (1) SCRIPTed 19.

14 Agencia Espa~nola de Protecci_ on de Datos Press Release, 'The AEPD Sanctions Google for Serious Violation of the Rights of the Citizens' (19 December 2013) <<https://googleblog.blogspot.ca/2012/01/updating-our-privacy-policies-and-terms.html>> accessed 30 September 2016.

15 For the purposes of this article, I use Google instead of Alphabet because of the familiarity of the name 'Google' for most people.

16 See AEPD (n 14) 1.

Emergence in data privacy: a holistic approach _ 143

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

introductions, Google and Facebook have acquired more than 150 and 50 companies each, respectively. Furthermore, companies combine data across these different processing activities based on separate purposes.¹⁷ In light of these developments, the focus on individual processing activity overlooks the fact that the totality of personal data collected based on separate purposes and/or the combination of data across these processing activities could be, as discussed in section 'Privacy interests in wholeness and sum', a source of concern for the individual (eg overexposure) or society at large (eg loss of practical obscurity).

Based on discussions about emergent property, this article calls for a holistic approach based on enhanced responsibility to certain entities, taking account of different factors including the totality of the processing activities and data aggregation practices. The holistic approach is akin to the regulation under competition law, where companies with a monopoly or dominant position are subject to special responsibilities. I maintain that data privacy will benefit from a similar tiered framework aimed at imposing enhanced responsibilities based on the holistic view of the totality of the processing activities and data aggregation practices. My aim is not to condemn the individualistic approach but to highlight instances where the current approach might benefit from a complementary approach based on a holistic view. At this juncture, it is important to highlight whether the newly adopted Regulation, which comes into force on 25 May 2018, changes the paradigm in relation to the above problem. The answer is both 'yes' and 'no'. It is 'no' because, despite the main motivation for the reform being to align the rules with new business practices,¹⁸ the core principles and the manner of their application (ie the individualistic nature) remain unchanged.¹⁹ However, the Regulation also introduces changes that give the regulatory authorities more tools for dealing with the identified problem including scalable penalties for breach and Data Protection Impact Assessment (DPIA). More particularly, the Regulation opens the door, albeit for more economic reasons, for conducting DPIA covering operations beyond 'a single project' (ie processing

activity).²⁰ Moreover, DPAs are entitled, subject to notification to the European Data Protection Board, to add processing operations that could be subject to DPIA.²¹ Thus, with an open eye for a holistic approach, these tools should give more flexibility to cater to the interests identified in this article.

17 Google's 2012 change of privacy policy is a good example. See Google Official Blog, 'Updating Our Privacy Policies and Terms of Service' (24 January 2012) <https://googleblog.blogspot.ca/2012/01/updating-our-privacy-policies-and-terms.html>, last accessed on 30 Sep 2016.

18 See Commission Communication, 'Safeguarding Privacy in a Connected World: A European Data Protection Framework for the 21st Century' COM (2012) 9 final 1.

19 See Peter Hustinx, 'EU Data Protection Law: The Review of Directive 95/46 EC and the Proposed General Data Protection Regulation' (European Data Protection Supervisor) 28 <https://secure.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/EDPS/Publications/Speeches/2014/14-09-15_Article_EUI_EN.pdf> accessed 30 Sep 2016 noting that 'in spite of all innovation – there is also a lot of continuity (in the Regulation). All the familiar basic concepts and principles will continue to exist, subject to some clarification and smaller changes in details' (emphasis in original). See also Orla Lynskey, 'The Foundations of EU Data Protection Law (OUP 2015) 2–3 (indicating that the main architecture of the rules remains the same).

20 GDPR, Rec 39.

21 GDPR, Art 35(4).

144 Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

Throughout this article, I use Google and Facebook as examples to demonstrate some of the challenges, but they are by no means the only entities engaged in such practices. Moreover, none of these examples are meant to deny the enormous social and economic benefits that those companies bring to users at (seemingly) no cost. The remainder of the article is structured as follows. First, I explain the concept of emergent property and how I intend to use it in relation to the data privacy discussions. Secondly, I use the concept for explaining the challenges posed by the commercialization of personal data and big data for data privacy. Third, I examine the privacy interests that emerge with such practices. Finally, I discuss the potential way forward in dealing with the identified challenges.

EMERGENT PROPERTY

The idea of emergence is developed as an alternative to scientific inquiry based on reduction—that is, reducing a phenomenon into its components and closely examining these constituent parts.²² Physics is a good example of a reductionist discipline. Although the scientific importance of reduction is not disputed, it is argued that reduction is less helpful for studying certain things, such as the behaviour of human groups (how a group behaves cannot necessarily be explained by the behaviours of the individual members). Thus, emergence fills this gap by studying the phenomenon as a collective whole.²³ In other words, while reduction explains Xs as 'nothing more than Ys', emergence explains Xs as 'something over and above Ys'.²⁴ Emergence is discussed in different forms and emergent property is one of them.²⁵ At its core, the concept of emergent property underlines that 'the whole is more than the sum of its parts', where the 'whole' represents the 'emergent property'. Despite its wider use across many fields,²⁶ there is no clear designation of emergent property as a theory or methodology or something else. In some instances, emergent property is discussed as a theory of its own,²⁷ while in other instances, it is discussed as part of a larger theory of complexity, whereas in certain instances, it is discussed as a methodology.²⁸ Another author indicates that emergent property is a concept discussed in a search for a theory.²⁹

22 Paul Davies, 'The Physics of Downward Causation' in Paul Davies and Philip Clayton (eds), *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion* (OUP 2006) 35.

23 Harold Morowitz, *The Emergence of Everything: How the World Became Complex* (OUP 2002) 14.

24 Robert Van Gulick, 'Reduction, Emergence and Other Recent Options on the Mind/Body Problem. A Philosophic Overview' (2001) 8 (9–10) *Journal of Consciousness Studies* 16.

25 Philip Clayton, 'The Re-emergence of Emergence: The Emergentist Hypothesis from Science to Religion' in Paul Davies and Philip Clayton (eds), *The Re-Emergence of Emergence: The Emergentist Hypothesis from Science to Religion* (OUP 2006) 2 (discussing four kinds of emergence).

26 Despite the lack of a clear definition, the concept of emergent property can be found across many domains, including biology, ecology, geophysics and social science disciplines, such as economics. See Eric Bonabeau, Jean-Louis Desselles and Alain Grumbach, 'Characterizing Emergent Phenomena (1): A Critical Review' (1995b) 9 (3) *Revue internationale de systématique* 336.

27 Morowitz (n 23) 10 (discussing the use of emergence theory in statistical mechanics).

28 *ibid* 1.

29 Peter Corning, 'The Re-emergence of "Emergence": A Venerable Concept in Search of a Theory' (2002) 7 (6) *Complexity* 20 (noting the difficulty of attaching a date 'to the re-emergence of emergence as a legitimate and mainstream concept').

Emergence in data privacy: a holistic approach _ 145

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

This notwithstanding, the concept of emergent property has been central to general systems thinking. In this context, the concept helps clarify the property of a system that is built upon more than one component. This thinking enables systems engineers to look beyond the properties of individual components in a system and understand the system as a result of a collective interaction among the components or as a single complex.³⁰ Complexity is particularly essential in characterizing the emergent property 'because emergence can be considered as an increase in the complexity of a system'.³¹

From a general systems perspective, a complex can take three forms.³² First, a complex can arise from the summation of its component parts (ie 'summative' complex). Secondly, it can emerge from the 'special characteristics' of the components.³³ Thirdly, the complex can result from the 'interrelations' of the components.³⁴ As shown in [Figure 2](#), these distinctions are important in drawing comparisons of how a complex emerges in data privacy law. The literature on emergent property varies as to what characterizes the emergent property, for example, in terms of its composition, structure or functionality.³⁵ Gulick distinguishes between three levels of emergence—namely, specific value, modest emergent properties and radical emergent properties.³⁶ In specific value emergence, the whole and the constituent parts share the same features, but the whole is different in type or value.³⁷ The summative complex falls under this category. Some argue that specific value emergence should not qualify as emergent property.³⁸ In modest kinds of emergence, the whole exhibits features that are different in kind from those of its constituent parts.³⁹ In radical emergence, the whole is both different in kind from the features of its constituent parts and its 'nature and existence is not necessitated by the features of its parts, their mode of combination . . .'.⁴⁰ There are questions concerning the existence of real cases of radical emergence.⁴¹ Thus, the most commonly discussed emergent property is emergence of the modest kind.⁴²

30 Misha Hebel, 'World-views as the Emergent Property of Human Value Systems' (1999) 16 (3) *Systems Research and Behavioral Science* 253.

31 Eric Bonabeau, Jean-Louis Dessalles and Alain Grumbach, 'Characterizing Emergent Phenomena (2): A Conceptual Framework' (1995a) 9 (3) *Revue internationale de systématique* 347.

32 Ion Georgiou, 'The Idea of Emergent Property' (2003) 54 (3) *Journal of the Operational Research Society* 240.

33 'For example, a complex may be constituted by one poodle, one Dalmatian and one Alsatian. This complex is understood as being of three elements belonging to the species "dogs"; it is thus a complex of dogs.' *ibid.*

34 'For example, a complex may be constituted by a palomino stallion, his mare and their pony. This complex is thus one of a family of palomino horses.' *ibid.*

35 Timothy O'Connor, 'Emergent Properties' (1994) 31 (2) *American Philosophical Quarterly* 91.

36 Van Gulick (n 24) 17.

37 *ibid.*

38 Jaegwon Kim, 'Being Realistic about Emergence' in Paul Davies and Philip Clayton (eds), *The Reemergence of Emergence: The Emergentist Hypothesis from Science to Religion* (OUP 2006) 191–92.

39 Van Gulick (n 24) 17.

40 *ibid.*

41 *ibid.* 18.

42 Kim (n 38) 192.

146 _ **Emergence in data privacy: a holistic approach**

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

It is beyond the purview of this article to delve into the details of such characterization.

This is because, first, the idea of emergent property is used as a lens for examining data privacy challenges; in other words, the main arguments put forth can, although examined through a systems thinking lens, stand independent of the concept of emergent property.⁴³ Furthermore, despite its origins in systems thinking, the idea of emergent property is also viewed as an epistemological theory which provides a foundation for construing knowledge of the whole in contexts other than systems.⁴⁴

For example, the idea of emergent property has been used to explain and understand the unforeseen consequences of interactions among different behaviours that form a system in its broader sense, such as a system of law.⁴⁵ A good example in this regard comes from laws that require automobile drivers to wear seatbelts while driving. The objective of such laws, not surprisingly, is to ensure the safety of drivers, mitigating injuries and deaths from accidents. This notwithstanding, at times, such laws can lead to a collective behaviour that contradicts the intended objective. In the example at hand, the compulsory seatbelt rules might not reduce the overall injuries or deaths from accidents because 'in wearing a seatbelt, drivers usually feel safer and willing to take greater risks'.⁴⁶ In this example, the property 'drivers usually feel safer and are willing to take greater risks' emerges, representing the emergent property. By regarding the system and its components as a single complex, the idea of emergent property helps to uncover hidden assumptions or unintended consequences of the collective interaction among the components or behaviours. This would enable the consideration of the potential collective behaviour and the more effective design of the system and its enforcement machinery. Similarly, the concept of emergent property could provide a useful lens for explaining and understanding some of the challenges that result from entities processing personal data for manifold purposes and data aggregation practices.

Thus, in this article, the concept of emergent property serves three related purposes. First, it helps to demonstrate why the individualistic approach is problematic in light of the growing commercialization of personal data and big data practices. Secondly, it serves as a lens for uncovering the privacy interests that emerge independent of and/or in combination with the individual processing activities. Thirdly, it provides useful guidance on constructing the potential way forward that will help address the identified challenges. The following sections are organized to address these issues in this order.

THE IDEA OF EMERGENT PROPERTY IN DATA PRIVACY

This section applies the discussions about emergent property for data privacy law. It starts off by briefly reviewing developments in case law from both sides of the Atlantic that reflect some emergent property reasoning in relation to government

43 That is, the concept is used as 'a way of seeing things which were previously overlooked or bypassed, and in this sense is a methodological maxim'. Ludwig Von Bertalanffy, 'The History and Status of General Systems Theory' (1972) 15 (4) *Academy of Management Journal* 424.

44 Georgiou (n 32) 243; see also Von Bertalanffy, *ibid* 423.

45 Georgiou, *ibid* 241.

46 *ibid*.

Emergence in data privacy: a holistic approach _ 147

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

processing of data. Then it identifies business practices in processing personal data that lead to the emergence of a property. This is followed by a discussion on how to understand emergence for data privacy purposes.

The emergence of emergent properties in case law

There are emerging decisions from the EU and US courts that reflect some emergent property thinking, although the discussions in the USA are a step ahead. In the EU, this line of reasoning is observed in the Court of Justice of the European Union (CJEU) judgment that invalidated the Data Retention Directive.⁴⁷ In that judgment, the Court emphasized the wide array of data about a private person that providers of electronic communication services or networks are required to retain, including data concerning 'the source of a communication and its destination, . . . the date, time, duration and type of a communication, . . . users' communication equipment, and . . . the location of mobile communication equipment, the name and address of the subscriber or registered user, the calling telephone number, the number called and an IP address for Internet services'.⁴⁸ According to the Court, '[t]hose data, taken as a whole, may allow very precise conclusions to be drawn concerning the private lives of the persons whose data has been retained . . .'.⁴⁹ For the Court, the collection and storage of such data constitutes interference with the right to respect for private life and the fundamental right to the protection of personal data, which can only be

derogated insofar as it is 'strictly necessary' in a democratic society.⁵⁰ Noting 'the vast quantity of data whose retention is required by that directive', the retention of that amount and kind of data fails to fulfil the strictly necessary requirement.⁵¹ The judgment, particularly the reference to 'taken as a whole', reflects some emergent property reasoning in the sense that although the retention of different sets of data individually might not constitute a breach of the rights of individuals, collectively, they constitute a breach of the right to respect for private life and the fundamental right to the protection of personal data. This implies that the whole exhibits features (ie breach of individual's data protection right) that do not exist in the constituent parts (individual data sets).

Across the Atlantic, the discussions are even more advanced. Following the 2012 decision of the US Supreme Court in *United States v Jones*,⁵² concepts similar to emergent property, such as the 'mosaic theory'⁵³ and 'quantitative

47 C-293/12 and C-594/12 *Digital Rights Ireland v Minister for Communications & Others* (2014) E.C.R. I-238 para 26.

48 *ibid*, para 26.

49 *ibid*, para 27 (emphasis added).

50 *ibid*, para 37–40.

51 *ibid*, para 66.

52 (2012) 132 Supreme Court 945.

53 Orin Kerr, 'Mosaic Theory of the Fourth Amendment' (2012) 111 *Michigan Law Review* 311; Christopher Slobogin, 'Making the Most of *United States v Jones* in a Surveillance Society: A Statutory Implementation of Mosaic Theory' (2012) 8 *Duke Journal of Constitutional Law & Public Policy* 12; David Gray and Danielle Citron, 'A Shattered Looking Glass: The Pitfalls and Potential of the Mosaic Theory of Fourth Amendment Privacy' (2013a) 14 (2) *North Carolina Journal of Law and Technology* 381; Monu Bedi, 'Social Networks, Government Surveillance, and the Fourth Amendment Mosaic Theory' (2014) 94 *Boston University Law Review* 1809.

148 _Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

privacy',⁵⁴ began attracting considerable attention in the US literature in relation to the Fourth Amendment.⁵⁵ The mosaic theory was initially articulated in *United States v Maynard* by the Court of Appeals for D.C. Circuit, which was later reviewed by the Supreme Court decision under the name *Jones*.⁵⁶ In that case, the police, investigating a drug conspiracy involving the defendant Mr Jones, installed a GPS without a warrant⁵⁷ on the defendant's car to track his movements over the course of one month.⁵⁸ Such tracking resulted in 2000 pages of data, which were then used to coordinate the defendant's movements and convict him (and Mr Maynard) of drug conspiracy.⁵⁹

Mr Jones appealed the decision to the D.C. Circuit Court on the grounds that the warrantless tracking via GPS had violated his 'reasonable expectation of privacy' under the Fourth Amendment.⁶⁰ One of the doctrines in assessing the right under the Fourth Amendment is the public disclosure doctrine which precludes a reasonable expectation of privacy of an individual's movements in public spaces. This means the mere fact that the information is exposed to the public view precludes protection under the Fourth Amendment (and the need for a warrant). Given that the GPS was tracking Mr Jones' movements on public roads, this would mean that Mr Jones had no reasonable expectation of privacy that would warrant protection under the Fourth Amendment.⁶¹ However, the D.C. Circuit Court disagreed and found the following:

[T]he totality of Jones's movements over the course of a month—was not exposed to the public: First, unlike one's movements during a single journey, the whole of one's movements over the course of a month is not actually exposed to the public because the likelihood anyone will observe all those movements is effectively nil. Second, the whole of one's movements is not exposed constructively even though each individual movement is exposed, because that whole reveals more—sometimes a great deal more—than does the sum of its parts.⁶²

The decision shows that although it might be legal for government to carry out a GPS tracking of single trips in public spaces, the aggregation of numerous trips over

54 David Gray and Danielle Citron, 'The Right to Quantitative Privacy' (2013b) 98 *Minnesota Law Review* 62; Danielle Citron and David Gray, 'Addressing the Harm of Total Surveillance: A Reply to Professor Neil Richards' (2012) 126 *Harvard Law Review Forum* 262.

55 'The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no warrants shall issue, but upon probable cause, supported by oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized': US Constitution Fourth Amendment.

56 (2010) 615 Court of Appeals, Dist. of Columbia Circuit 544.

57 In fact, the police had obtained a warrant to install GPS, but violated the conditions for installing it within 10 days and within the district of Washington. See Gray and Citron (n 53) 392.

58 United States v Maynard (n 56) 948.

59 *ibid.*

60 *ibid.*

61 The lower court accepted Mr Jones' challenge to track the location of the vehicle on Mr Jones's private premises without a warrant, but considered the tracking on public roads as lawful citing the Supreme Court decision in United States v Knotts (1983) 460 Supreme Court 277.

62 United States v Maynard (n 56) 558 (emphasis added).

Emergence in data privacy: a holistic approach _ 149

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

the course of a month becomes illegal because it reveals more private information than the individual constituent parts (individual trips or locations). In other words, individuals may not have a reasonable expectation of privacy for their single public movements, but an aggregation of these single public movements could create a reasonable expectation of privacy. The decision of the Appellate Court was reviewed by the Supreme Court on certiorari and resolved on the grounds of physical trespass due to the installation of the GPS on private property (ie car).⁶³ Although the unconstitutionality of the mere installation precluded the Supreme Court to entertain the issue of 'aggregation' in greater detail, Justices Alito and Sotomayor concurred with the D.C. Circuit Court's stance that citizens could invoke a reasonable expectation of privacy regarding aggregated data—ie the whole—despite the lack thereof in the individual trips that constitute the whole.⁶⁴ Both the judgments from the Circuit Court and the concurrences of the US Supreme Court recognize the emergence of a property (ie reasonable expectation of privacy) despite the absence of such a feature in the constituent parts.

In this article, the goal is to examine the extent to which such concepts of the mosaic theory, or wholeness could be applicable to the EU data privacy framework, not only in light of government practices but also regarding the processing of personal data by private entities. However, in this article, the concept of emergent property is preferred over the mosaic theory and quantitative privacy. This is because the use of the mosaic theory and quantitative privacy has been limited to Fourth Amendment discussions in the context of government surveillance and the collection of data in public spaces, which in turn limits their application to data processing by private entities. The core question behind the mosaic theory and quantitative privacy discussions is whether the aggregation of data from public spaces creates a 'reasonable expectation of privacy' and constitutes a 'search' under the Fourth Amendment.⁶⁵ It seems that the main difference between the mosaic theory and quantitative privacy is that the former focuses on 'how much' information is collected, whereas the latter focuses more on 'the technology used' to collect the information.⁶⁶ Given the variety of techniques and technologies used by private entities to collect data and the challenges of measuring the amount of information collected, these concepts are not well-suited to processing by private entities.⁶⁷ More importantly, this article focuses on data processing practices, and particularly the processing of personal data for multiple purposes and the data aggregation practices across the different processing operations. In this sense, the concept of emergent property helps to better explain how and what privacy interests can emerge independent of and/or in combination with the particular interests in the individual processing activities when an entity processes personal

63 See United States v Jones (n 52) 949.

64 Citron and Gray (n 54) 270. Justice Alito reasoned that short-term tracking of a person's public movements is not protected under the Fourth Amendment, but a longer term could be. United States v Jones, *ibid* 964 (Alito J, concurring).

65 In fact, the mosaic theory has its origin in the financial sector and is used in practices of aggregation of publicly available information. Its usage in relation to the Fourth Amendment reflects this origin in the sense that the discussions are focused on the aggregation of information collected from public spaces.

66 According to the proponents of the concept, quantitative privacy is triggered if 'the technology facilitates broad and indefinite surveillance'. In this sense, it is technology-specific. See Gray and Citron (n 54) 72.

67 See the limitations of the mosaic theory in Kerr (n 53); see also Bedi (n 53).

data for multitudinous purposes. The concept also provides useful guidance in terms of the potential solutions that need to be considered in addressing the identified challenges. Next, the article identifies the emerging data processing practices that are common across different businesses that lead to the emergence of property.

Data processing practices leading to emergence

Key considerations in driving the EU and the US courts towards a holistic view are the volume and variety of data and the aggregation of such data. In the same fashion, the motivation for using the concept of emergent property in data privacy law could be attributed to two emerging data processing practices. The first relates to the growth in the number of entities that process individuals' personal data for a panoply of purposes. The increasing value of personal data for commercial purposes creates the desire to collect and amass as much personal data both through legitimate and illegitimate mechanisms, leading to the emergence of companies processing vast amounts of personal data for a wide array of purposes. The second, but related, development pertains to big data practices that represent substantial value in the repurposing, recombining and multi-purposing of personal databases. These developments represent the emergence of a property—that is, complexity—which is overlooked under current data privacy rules. As discussed further below, these respective developments can be compared to the summative and interrelation aspects of a complex of emergent properties.

In data privacy, the summative aspect as a complex emerges when entities engage in multiple processing activities based on separate purposes. In the digital economy, the volume and quality of the personal data controlled by companies is becoming a key source of revenue and market power. References to personal data as 'the new currency', 'the new oil' and the 'life blood of businesses' are tributes to its paramount importance in the digital economy. At the heart of the business model for companies such as Google or Facebook, is a detailed collection and analysis of consumer data, often gathered without the individual's knowledge or consent. Such information is used to target advertisements to specific groups whose members might be most interested in buying certain products or services. Moreover, personal data can be bought, sold and traded on its own.⁶⁸

Such growing importance of personal data for commercial purposes, coupled with the ever-sinking cost of storage, is creating a drive for a 'digital land grab', shifting the priorities of organizations to collect and harness as much personal data as possible in order to maximize their market position.⁶⁹ This desire is being pursued using legitimate grounds, such as expanding to new product sectors and acquiring entities with valuable personal information. Google's aggressive expansion to a broad array of new product areas is a case in point. At present, Google offers more than 100 services, of which more than 70 are offered for free. This number is not only a result of legitimate acquisitions, but also controversial practices of nudging users to consume

⁶⁸ William Eggers, Rob Hamill and Abed Ali, 'Data as the New Currency: Government's Role in Facilitating the Exchange' (2013) 13 Deloitte Review 19.

⁶⁹ White House, 'Big Data: Seizing Opportunities, Preserving Values' (Executive Office of the President 2014) 54.

the new product or subscribe to the service launched. When launching its social networks, individuals with a Gmail account were automatically given a Buzz account (now defunct)⁷⁰ and later a Google+ account.⁷¹ The end result is a large pool of processing activities involving data that are collected legally and controversially.

This development gives rise to a complexity that does not exist when entities process personal data for a single or handful of purposes. This can be likened to the 'summative' complexity in emergent property, where the 'whole can be understood as the summation of processing activities based on separate purposes' ('sum' hereinafter). The summative aspect is possibly one that gives rise to scepticism from proponents

of the emergent property mainly because the emergence is quantitative (value) rather than kind (qualitative).⁷² However, this is not always the case. In the Article 'More is Different', Philip Anderson, a Nobel laureate and a person credited for his significant contribution to the writings on emergent property,⁷³ indicates that a change in scale or quantity often transforms into qualitative change in the behaviour of a system.⁷⁴ According to him, the summation of less complicated pieces into a system can result in a new type of behaviour that raises new questions.⁷⁵ As shown in the next section, this holds true for data privacy because there are privacy interests independent of the individual processing when an entity processes personal data for a number of purposes (regardless of an 'actual' aggregation).⁷⁶ From a data privacy point of view, this means that processing data for a wide array of purposes paints a different risk picture than processing for a single or handful of purposes. In this sense, from a data privacy perspective, the sum has features (risk) that are different in kind (qualitative change) from the features in the individual processing activities—and thus can be considered to meet the requirements of the modest kind of emergence. As noted in section 'Emergent property', an emergent property is characterized as the modest kind if the whole exhibits features that are different in kind from those of its constituent parts.⁷⁷ Furthermore, the potential for aggregation that exists leads to the emergence of new property as a result of such aggregation or combination of data across different processing activities. This property (complexity) that emerges due to data aggregation practices can be likened to the interrelation aspect.

70 In this case, the FTC found this practice to be deceptive and Google subsequently agreed to implement a comprehensive privacy plan including independent privacy audits for the following 20 years. FTC Press Release, 'FTC Charges Deceptive Privacy Practices in Google's Rollout of Its Buzz Social Network' (30 March 2011) <<https://www.ftc.gov/news-events/press-releases/2011/03/ftc-charges-deceptive-privacy-practices-googles-rollout-its-buzz>> accessed 30 Dec 2016.

71 See Alexei Oreskovic, 'Google Linking of Social Network Contacts to Email Raises Concerns' (Reuters, 9 January 2014) <<http://www.reuters.com/article/us-google-gmail-idUSBREA081NH20140110>> accessed 30 Sep 2016.

72 Kim (n 38) 191–92.

73 See the introduction written by Goldstein in republishing Anderson's work. Jeffrey Goldstein and Philip Anderson, 'More Is Different: Broken Symmetry and the Nature of the Hierarchical Structure of Science' (2014) 16 E:CO 117.

74 Philip Anderson, 'More Is Different: Broken Symmetry and the Nature of the Hierarchical Structure of Science' (1972) 177 Science 393.

75 *ibid* 396.

76 Solove defines aggregation as gathering and combining information about a person; see Solove (n 1) 118.

77 Van Gulick (n 24) 17.

152 _ Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

The growing tendency to aggregate or combine data is associated with 'big data' practices. What is to be defined and described as big data is a question that goes beyond the purview of this article. For the purpose of this article, it suffices to point out the changes in the volume and variety of data.⁷⁸ Equally important, nonetheless, is the societal change that accompanies big data practices. At the centre of such change is society's shift from the static view of data, where data are collected for a singular purpose or one-off use, to innovative and novel secondary uses.⁷⁹ More importantly, the underlying logic with big data practices, as Mayer-Schoenberger and Cukier put it, is that 'the sum is more valuable than its parts, and when we recombine the sums of multiple datasets together, that sum too is worth more than its individual ingredients'.

⁸⁰ Thus, big data practices represent substantial value in the repurposing, recombining and multi-purposing of personal databases. Such a practice of collecting data for multiple purposes, together with combining such data across data sets, represents another complexity that does not exist when the processing of personal data is done in silos. This complexity is comparable to the interrelation aspect in discussions about emergent property, where the whole can be understood as the 'interrelation of data sets collected for separate purposes' ('wholeness' hereinafter). The interrelation aspect signifies the need for some kind of interaction among the data sets collected based on separate purposes.⁸¹

This aspect of emergent property—ie wholeness—as an issue was partly witnessed

when Google changed its privacy policy in 2012. Up until 2012, Google had collected individuals' personal data in offering different services (under different accounts)—that is, for separate purposes. By March of the same year, Google had taken the initiative to consolidate all of its services under one account. Before the consolidation, Google, at least in theory, claimed that a user's data from, for example, YouTube, Gmail and Search were kept separately. However, following the consolidation, Google was able to integrate the user's data from YouTube, Gmail and virtually all other services to create a more comprehensive profile of the user. This profile is akin to the concept of emergent property in the sense that it contains more value for Google than the individual profiles under the separate Google accounts. It also represents a risk that might not exist with the individual accounts kept separately. At least in Europe, different stakeholders including the Article 29 Working Party have voiced

78 According to the National Institute of Standards and Technology (NIST), 'Big Data consists of extensive datasets primarily in the characteristics of volume, variety, velocity, and/or variability that require a scalable architecture for efficient storage, manipulation, and analysis'. See NIST, 'NIST Big Data Interoperability Framework: Volume 1, Definitions' in Nancy Grady and Wo Chang (eds), NIST Special Publication 1500-1 (Final Version 1, 2015) (emphasis added).

79 Viktor Mayer-Scho'nberger and Kenneth Cukier, *Big Data: A Revolution that Will Transform How We Live, Work, and Think* (Houghton Mifflin Harcourt 2013) 96.

80 *ibid* 108 (emphasis added).

81 More generally, from an information science perspective, the interrelation can be approached in what is referred to as mutual information: 'Mutual information quantifies the information processes going on between two elements A and B: intuitively, there is mutual information between A and B if A and B are able to affect one another's behavior.' Bonabeau, Dessalles and Grumbach (n 31) 364. In this sense, an interrelation is considered to exist between two data sets collected for different purposes if new information, unbeknownst to the two data sets, can be derived from them or the data sets affect each other in any other way.

Emergence in data privacy: a holistic approach _ 153

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

their concern on the matter. Some of these concerns are discussed in section 'Privacy interests in wholeness and sum'.

In light of the above discussions, the idea of emergent property from a data privacy point of view can be approached from two related angles: first, wholeness, which represents a number of processing activities based on separate purposes with interactive elements among them, such as aggregation; and secondly, the summative aspect, which represents the scope of the processing activities in which an entity is engaged, as measured by the sum, based on separate purposes without an actual interactive element among them, but with the potential to do so. Figure 2 above depicts the conceptual distinction.⁸²

More generally, for the purposes of data privacy, emergent property can be approached as a way of looking at multiple processing activities (n) as a single complex, where an entity aggregates data across the processing activities or with the potential to do so, and a means for examining the privacy interests that emerge in such a complex independent of, or in combination with, privacy interests in the individual processing activities. The next question is 'What are the privacy interests in the wholeness and sum?' More particularly, what privacy interests are there, for example, if Google processes personal data through 100 separate purposes (n/4100) so far as each processing activity is compliant? The following section undertakes that task.

PRIVACY INTERESTS IN WHOLENESS AND SUM

The term 'privacy interest' is used to denote the concerns or values that data privacy rules aim (or ought) to safeguard.⁸³ The concerns identified here are not necessarily novel. Some of them could be viewed as not fully safeguarded by the current rules; thus, they would need to be considered going forward. Others are existing concerns

Processing
2
Processing
n
Processing
1
Processing
2
Processing
n
Processing
1

Summative Interrelation (wholeness)

Figure 2. Summative and interrelation (wholeness) aspects.

82 This notwithstanding, the difference between the ideas of wholeness and summative is not significant because it is common practice for entities to aggregate data across processing activities or there exists the potential to do so. For this reason, they are treated as wholeness and sum.

83 For more, see Bygrave (n 5) ch 7.

154 Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

that become prominent or are exacerbated by the emerging data processing practices discussed above. For this reason, an attempt has been made to link the concerns to these data processing practices.

It is tempting to posit that the larger the wholeness and sum becomes (number of processing activities based on separate purposes), the less individual privacy there is. This is even truer when there is a combination of data across the different processing activities. The retort to this argument is that if each processing is compliant with data privacy rules, there should be no need for concern about the wholeness and sum of the processing activities based on a number of distinct purposes. This is a way of arguing that since all the parts of a system are compliant, so is the system as a collective whole. In fact a dissenting judge in the D.C. Circuit Court judgment argued to the same effect, indicating that '[t]he sum of an infinite number of zero value parts is also zero.' ⁸⁴ Put differently, the legitimacy of each activity (the constituent part of the whole) legitimizes the plurality (the whole). In the discussions about emergent property, this argument is referred to as the 'fallacy of composition', ⁸⁵ where one attributes a property of the component part of a system (compliance in this case) to the system as a whole. More importantly, my argument is not that wholeness and sum violates data privacy principles outright and in all circumstances; instead, the argument is that there are privacy interests that necessitate some regulatory oversight independent of, and/or in combination with, the individual processing activity, particularly where companies process personal data through a wide array of services or products, where almost every use of the service generates personal data, and the data are combined across these processing activities. In what follows, I highlight these privacy interests.

Protection against overexposure of individuals

This section examines how the practices of processing personal data for multiple purposes together with data aggregation across processing activities overexpose individuals, thereby undermining their personal autonomy, integrity and dignity. As Bygrave shows, these are interests for data privacy rules to safeguard.⁸⁶ More specifically, the interest in protection against overexposure is closely related to what Bygrave refers to as Group 2 interests, particularly 'non-transparency (i.e. a person's interest in avoiding being rendered transparent vis-a-vis other persons and organizations)'.⁸⁷

The focus of the current rules on the 'individual processing' activity means that they overlook the exposure and transparency of the individual resulting from the totality of the data collected by entering a wide array of new product areas, acquisitions and combinations of data across different processing operations. Again, with nearly 100 processing activities at its disposal, Google serves as a good example.

Apart from the personal information collected during product or service registration, as part of its YouTube service, Google collects data on the types of videos a

⁸⁴ United States v Jones, 625 F 3d 766, 769 (DC Cir 2010) (Sentelle J, dissenting).

⁸⁵ William L Rowe, 'The Fallacy of Composition' (1962) 71 (281) Mind 87.

⁸⁶ See Bygrave (n 5) ch 7 on 'Values and Interests Safeguarded by Data Protection Law'.

⁸⁷ *ibid* 150.

Emergence in data privacy: a holistic approach _ 155

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

user watches and the user's likes and dislikes. As part of its Search, Google gets insights on what products a user is interested in, the books he/she might want to purchase and an illness he/she might have. Google knows our contact list from Gmail, Google+ and Groups. Google knows our gathering places from Calendar and Maps. Our locations are further revealed through Maps, Search and Earth, and Android

GPS tracks our location even when no apps are running. Credit card and bank information are accessible if one uses CheckOut, Finance or Google Wallet. Search, Gmail, Books and Health (before its discontinuation) might contain health information. Talk, Voice, Maps and Calendar signal destination plans. Our browsing habits are monitored and recorded through Google Chrome. Even more, through GoogleX, Google is becoming an Internet service provider (and also Facebook) offering Wi-Fi, wireless broadband (Fi) and fibre, providing a window into virtually everything a user does on the Internet. What we write (and think) is scanned from Gmail attachments, Google Drive and Dropbox. Millions of mobiles run on Google Android. Moreover, through its Nest Cam, Google is promising to 'help you look after your home and family – even when you're away. With 24/7 live streaming, advanced Night Vision, activity alerts, one app for all your Nest products, and a versatile magnetic stand, Nest Cam helps you keep an eye on what matters'.⁸⁸ Through its Analytics service, Google collects the browsing habits of individuals from millions of third-party websites. The list goes on; there is hardly anything that Google does not do or plan to do.⁸⁹ Organizations also resort to controversial means of collecting user data that provides valuable insights about consumers. The Google Street View Project is a good example, where Google accessed the communication of individuals transmitted over Wi-Fi networks in many countries, which also led to fines and confiscation of properties, for example, in North Korea.⁹⁰ It is claimed that the access was deliberately designed, as the resulting data are considered instrumental for the success of Google Maps and self-driving cars.⁹¹ Essentially, Google is our browser, our search engine, our messenger, our guide for driving and taking public transport and a platform for writing and storing our files. Our mobile phones run on Google. It will not be long before Google cars flood the streets. Even our houses are coming under Google's surveillance cameras. This means that even the usage of a fraction of these services creates a transparent data subject, an individual who exposes significant aspects of his/her life willingly in order to use the services. In using the services, data relating to the usage of the services, clicks and even cursor hovers over links are monitored, recorded, analysed and

⁸⁸ Nest Cam website <<https://nest.com/ie/camera/meet-nest-cam/>>

⁸⁹ This includes smart fabric or clothing, contact lenses and watches. See Zach Miners, 'Google's Project Jacquard to Make Smart Fabric for Smart Clothing, Levi's First Official Partner' (PC World, 29 May 2015) <<http://www.pcworld.com/article/2928372/this-smart-fabric-from-google-can-change-the-music-and-turn-off-the-lights.html>> accessed 30 Sep 2016.

⁹⁰ See Mark Burdon and Alissa McKillop, 'The Google Street View Wi-Fi Scandal and Its Repercussions for Privacy Regulation' (2014) 39 Monash University Law Review 702.

⁹¹ See Yasha Levine, 'Google's For-Profit Surveillance Problem' PANDO (16 December 2013) <https://pando.com/2013/12/16/googles-for-profit-surveillance-problem/> accessed 30 Dec 2016 arguing that '[i]t was all part of the original program design: Google had equipped its Street View cars with surveillance gear designed to intercept and vacuum up all the wireless network communication data that crossed their path.'

156 Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

combined, often with no or little knowledge on the part of the user. By so doing, as David Lyon comments, big data practices augment surveillance because 'it renders ordinary everyday lives increasingly transparent to large organizations'.⁹² Omer goes even further and observes that this is, in fact, surveillance of an omnipresent and indiscriminate nature, and there is no reason to treat it differently.⁹³ One expert describes Google's main business as 'for-profit intelligence', conducting an 'espionage operation on a global scale'.⁹⁴ Even Google itself is not shy to acknowledge its pervasive surveillance of individuals. Google's former CEO, Eric Schmidt, has been quoted as saying 'We know where you are. We know where you've been. We can more or less know what you're thinking about'. This is not limited to Google. Facebook also follows every digital move and accumulates troves of personal data.⁹⁵ What Facebook is able to collect through its website could easily fall under several service or product categories. In comparison to Google, Facebook seems to have mastered the ability to solicit a wide array of data using simple features and making it look like all the data being collected fall under the umbrella of a single service—thereby justifying its processing using only one specific (and legitimate) purpose.

Comparing the NSA's surveillance with what private entities such as Google and Facebook are doing, one author comments that the 'for-profit surveillance' by Silicon Valley 'dwarfs anything being run by the NSA'.⁹⁶ In fact, one important lesson from the recent revelations is that the surveillance by private entities is the engine that powers government surveillance. Thus, there seems to be consensus that what is happening is a wide ranging surveillance of individuals by private entities and governments in tandem and with an absence of clear boundaries to draw. Having quoted the definition of surveillance as 'systematic observation', Schneier argues that 'modern-day electronic surveillance is exactly that'.⁹⁷ The dangers of surveillance are widely discussed elsewhere and need not be reiterated here.⁹⁸ In fact, whether such reality should constitute 'surveillance' is less important than the fact that what is happening is an 'overexposure' of individuals to private entities (and governments).

A central objective of the data privacy rules is to allow the individual to selectively self-disclose his/her information based on free choice, referred to as information self-

⁹² David Lyon, 'Surveillance, Snowden, and Big Data: Capacities, Consequences, Critique' (2014) 1 (2) Big Data & Society 4.

⁹³ Omer Tene, 'What Google Knows: Privacy and Internet Search Engines' (2008) 1433 Utah Law Review 40.

⁹⁴ Yasha Levine, 'The Psychological Dark Side of Gmail: Google Is Using Its Popular Gmail Service to Build Profiles on the Hundreds of Millions of People Who Use It' AlterNet (31 December 2013).

⁹⁵ Bernard Harcourt, *Exposed: Desire and Disobedience in the Digital Age* (Harvard UP 2015) 454–55 (Kindle locations). Citing Eben Moglen, Harcourt notes that '[i]f you have a Facebook account, Facebook is surveilling every single moment you spend there'.

⁹⁶ Levine (n 94).

⁹⁷ Bruce Schneier, *Data and Goliath: The Hidden Battles to Collect Your Data and Control Your World* (WW Norton & Company 2015) 4.

⁹⁸ See, among others, Neil Richards, 'The Dangers of Surveillance' (2013) 126 (7) Harvard Law Review 1934; William Staples, *The Culture of Surveillance: Discipline and Social Control in the United States* (St. Martin's Press 1997); Paul Schwartz, 'Privacy and Democracy in Cyberspace' (1999) 52 Vanderbilt Law Review 1607; Reginald Whitaker, *The End of Privacy: How Total Surveillance Is Becoming a Reality* (The New Press 1999).

Emergence in data privacy: a holistic approach _ 157

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

determination, thereby preventing undue interference in his/her autonomy, integrity and dignity. This ensures the protection of the individual from 'manipulation or control by others'.⁹⁹ However, the more exposed an individual becomes, the easier it is to 'force his obedience' and suppress his capacity to make free choices.¹⁰⁰ Such limitless knowledge of the individual transforms into a significant power imbalance, where entities (private and government) become 'omnipotent parents and the rest of society . . . helpless children'.¹⁰¹ Harcourt echoes similar views claiming that overexposure renders individuals open and accessible to serve idiosyncratic corporate and governmental interests.¹⁰² David Lyon adds that the unprecedented access to individuals' data gives entities (private and government) the ability to define 'who they are, what they should desire or hope for, including who they should become'.¹⁰³ Overall, overexposure makes individuals powerless, turning them into predictable citizen-consumers' who can easily be stimulated and nudged to serve profitmaximizing goals.¹⁰⁴ This reality undermines the very essence of an individual's autonomy, integrity and dignity that data privacy rules, at least in the EU context, are meant to safeguard.¹⁰⁵

In this regard, it is important to note the link between the scope of the processing activity in which an entity is involved and the resulting exposure of individuals. The greater the number of processing activities due to the wide array of services or products (ie more economy of scope), the more exposed the individuals become, particularly if there is a possibility of combining data across these processing activities, and the more difficult it becomes for the individuals to control self-disclosure and exercise free choice. For example, one study claims that firms that can offer diverse services (ie more economy of scope) find it easier to convince consumers to consent than new entrants and SMEs.¹⁰⁶ It could be argued that the more services consumers use from an entity, the more they come to trust the company, and thus the company can easily obtain consent. However, this claim bears little relation to reality. According to a Eurobarometer survey, 63 per cent of Europeans do not trust online

businesses, like search engines.¹⁰⁷ Thus, an alternative explanation could be that when consumers use diverse services from a single entity, they become so dependent

99 Bygrave (n 5) 133–34.

100 Paul Schwartz and Joel Reidenberg, 'Data Privacy Law' (Michie 1996) 39; see also Daniel Solove, 'Privacy and Power: Computer Databases and Metaphors for Information Privacy' (2001) 53 Stanford Law Review 1396.

101 Schwartz and Reidenberg, *ibid* 39; see also Solove, *ibid* 1396.

102 Harcourt (n 95) 289–92 and 482–86 (noting that the more exposed we are to others, the more we lose 'our sense of control over our destiny and self-confidence, our sense of self').

103 Lyon (n 92) 7.

104 Julie Cohen, 'What Privacy Is for' (2012) 126 Harvard Law Review 1919.

105 See discussions in Bygrave (n 5) 125ff. However, the danger of overexposure goes beyond just privacy interests. For example, Julie Cohen examines the effects of overexposure on self-development and innovation and concludes that overexposure and transparency can impair one's ability to freely tinker and experiment, which are essential for innovation and self-development; *ibid*.

106 James Campbell, Avi Goldfarb and Catherine Tucker, 'Privacy Regulation and Market Structure' (2015) 24 (1) Journal of Economics & Management Strategy 47.

107 Special Eurobarometer 431, Data Protection (European Commission 2015) 63.

158 *Emergence in data privacy: a holistic approach*

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

that they develop a sense of helplessness to refuse to provide consent.¹⁰⁸ Even more important is the claim by Schwartz and Reidenberg that the more the company knows the individual, the more ability it has to force the individual's obedience and suppress free choice.¹⁰⁹

Experiments by Facebook demonstrate the power that such private entities have gained to impose social control over their users. Facebook triggered outcries for manipulating users' behaviours through two separate experiments conducted on its users.¹¹⁰ Beyond the legitimacy of such experiments without users' knowledge,¹¹¹ these experiments have shown how easily the social network, with a deep knowledge of individuals' personal details, can influence users' moods and even their voting behaviours.

¹¹² The first example suffices for our purpose. On the basis of some algorithm, Facebook classified the moods conveyed by users' updates. Then a user or a set of users was shown updates on their Newsfeed¹¹³ conveying a specific mood. In order to do this, Facebook had to algorithmically suppress updates that did not convey the relevant emotion. In a sense, Facebook had to decide for its users which topics they should see and which topics should be hidden from their Newsfeed. By so doing, Facebook was able to influence the users' moods, meaning that the more the users were exposed to updates conveying a positive mood, the more positive posts (and the fewer negative posts) these users produced, and vice versa.¹¹⁴ Similarly, omitting both positive and negative emotional content from the user's Newsfeed reduced the number of words the person subsequently produced.¹¹⁵ One could just look at the Facebook experiment as a well-intended study aimed at contributing to academic discussion. However, a closer look reveals the commercial interest in learning about the causal effect of one friend's mood (or a set of friends' moods) on another person's mood (or set of persons' moods) or, more generally, how a person behaves at different times when exposed to the different moods of a friend or set of friends. Such knowledge would have enormous benefits in terms of the ads that Facebook decides to show to the user according to his/her friends' actions or behaviours.¹¹⁶ Recently, Facebook added a new feature that allows users to

¹⁰⁸ *ibid* 7 (noting that a large majority of Europeans [71%] 'say that providing personal information is an increasing part of modern life and accept that there is no alternative other than to provide it if they want to obtain products of services').

¹⁰⁹ Schwartz and Reidenberg (n 100) 39.

¹¹⁰ Adam Kramer, Jamie Guillory and Jeffrey Hancock, 'Experimental Evidence of Massive-Scale Emotional Contagion through Social Networks' (2014) 111 (24) Proceedings of the National Academy of Sciences 8788.

¹¹¹ Interestingly, in their study on emotional contagion, the authors claim that users agreed to such experiments when they expressed their consent to use the services. See *ibid*. The authors later apologized for the possible misunderstanding and anxiety created due to the research results. See BBC Technology, 'Facebook Emotion Experiment Sparks Criticism' (BBC, 30 June 2014) <<http://www.bbc.com/news/technology-28051930>>

¹¹² Robert Bond and others, 'A 61-Million-Person Experiment in Social Influence and Political Mobilization' (2012) 489 (7415) Nature 295.

¹¹³ Newsfeed is the feature that shows story updates (posts, photos, likes) from friends or pages to which the user is connected.

114 Kramer and others (n 110) 8788.

115 *ibid* 8789.

116 For example, it is common practice for Facebook to show products or services on the basis of friends' 'likes'.

Emergence in data privacy: a holistic approach _ 159

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

prioritize posts from family and friends. The announcement of the feature was accompanied by the following statement: 'Put the people you love at the top: We care about showing you posts from people who matter to you. We've made new controls that allow you to prioritize friends and family in your Newsfeed.' One could not help but wonder if this new feature had anything to do with the above-described experiment.

There is no evidence to suggest that the addition of this feature is a follow-up to Facebook's earlier study on users' moods, but one suspects that this is a well-thought-out plan from Facebook to know whom users are most sensitive to and how to better influence them.¹¹⁷ Facebook seems to have determined that the better it knows the individual, the more power it has to persuade that individual.¹¹⁸

This notwithstanding, in following the 'individualistic approach', the assessment of compliance under the EU data privacy rules continues to focus on each individual processing regardless of the scope and total number of processing activities an entity is conducting. Taking Google as an example, the focus on individual processing activity means that compliance with the data privacy rules, for example, in relation to Gmail, will be assessed on its own, as will the processing of data in relation to Google+, YouTube, Search and the remaining (close to 100) Google services. At least that would have been the approach before Google decided to consolidate the services under one account.¹¹⁹ This, however, overlooks the overexposure that emerges when entities are involved in multiple processing activities, where almost every use of the service generates personal data, and the data are combined across these processing activities. This means that data privacy rules need to look beyond the individual processing activities based on a specific purpose and relate to the totality of the data collected by entering a wide array of new product areas, acquisitions or combination of data from different sources. I believe that the holistic approach, which I discuss in section 'Towards a Holistic Approach: Enhanced Responsibility?', caters to such a need.

Loss of transparency and accountability

Generally, the discussion above on the wholeness and sum overexposes the individual; at the same time, it obscures the transparency and accountability of the entities in control of such data. More specifically, the existing accountability and transparency mechanisms based on the purpose limitation principle start to break down when the entities engage in combining data across different services based on

¹¹⁷ Often, much of the discussion on such influence is associated with the algorithms. However, that is only partially true. Mayer-Schönberger and Cukier convincingly argue that the primary drive of 'big data' practices is less about the changes in algorithms and more about access to 'more data'. According to the authors, a computer's ability to beat a human being at chess is not a result of an improvement in the algorithm behind the game but largely because the computer has access to 'more data' about possible moves and human behaviour; See n 79.

¹¹⁸ Richards (n 98) 1955 (noting that increased access to an individual's information gives 'increased power to persuade' the individual)

¹¹⁹ Or the alternative is to try and find a legitimate and specific purpose under the umbrella of generic 'Google services', which essentially makes, as is discussed in the next section, the purpose specification and limitation principle, and other principles that are built on it, ineffective.

160 _ Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

separate purposes.¹²⁰ The purpose limitation principle is of fundamental importance in the general application of data privacy rules and compliance with these rules.¹²¹ This is mainly because, as shown in section 'Individualistic approach to data privacy', the principle is an essential prerequisite in the application of the core data privacy principles under Article 6(1) of the Directive. Both transparency and accountability are central elements of the purpose limitation principle because having a specific and explicit purpose for processing enables data controllers to be transparent about how they use personal data both to regulators and data subjects, thereby ensuring the accountability

of entities processing personal data.¹²² Transparency is related to the openness of the entities processing personal data with both data subjects and regulators on what data are collected, for what purpose and how they are used.¹²³ Furthermore, accountability is related to the responsibility of the parties, and particularly the controller, to implement and demonstrate compliance with the data privacy principles enshrined under the Directive. In this sense, accountability implies an obligation (i) to adopt concrete and practical measures for the implementation of the core data privacy principles and to do so in a verifiable manner; (ii) to ensure that these measures are effective; and (iii) to assume liability when the relevant measures are not taken or are ineffective. However, the aggregation of data across different services or products obscures the transparency in terms of what data have been collected, for which purpose(s) and whether the data are necessary for that (those) particular purpose(s). Similarly, assessing whether a certain reuse of data is or is not compatible with original purposes becomes very challenging. The following paragraphs discuss these challenges in light of the two notions of the purpose limitation principle, namely purpose specification and compatibility.

Purpose specification

The notion of purpose specification comprises three aspects—the purpose must be ‘specific’, ‘explicit’ and ‘legitimate’. The ‘specificity’ aspect underlines the need for a precise and sufficiently defined purpose for the collection of personal data that must be determined prior to, or at the latest during, the collection of data.¹²⁴ The ‘explicit’ requirement is aimed at ensuring that the purpose is expressed unambiguously and is communicated in an intelligible manner, to ensure that all relevant stakeholders understand. The ‘legitimacy’ aspect requires a legal basis for processing personal data as provided under Articles 7 and 8 of the Directive. Among other things, the goal of

120 None of this is to deny the benefits of synchronization and the smooth experience one would have when using one account across different services, including reduced costs of reading privacy policies. See Aleecia McDonald and Lorrie Cranor, ‘The Cost of Reading Privacy Policies’ (2008) 4 Information System: A Journal of Law and Policy for the Information Society 543. However, data aggregation and policy consolidation are not mutually exclusive. Moreover, comprehensiveness more often than not comes at the cost of specificity.

121 WP203 (n 4) 18. The principle is described as the ‘cornerstone of data protection’ 4; see also Lee Bygrave, *Data Privacy Law: An International Perspective* (OUP 2014) 153.

122 WP203 (n 4) 18.

123 See Art 12 and Rec 38 of the Directive.

124 Rec 28 of the Directive states that the purpose ‘must be determined at the time of collection of the data’.

Emergence in data privacy: a holistic approach _ 161

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971> by University of Oslo Library. Library of Medicine and Health Sciences user on 21 February 2018

having a specific, explicit and legitimate purpose is to ensure the transparency and accountability of the controller to both regulators and data subjects on how the data are used.¹²⁵ This, in turn, ensures legal certainty and predictability.¹²⁶ Additionally, the purpose specification enables users to make informed decisions about their data, and thus is essential in strengthening users’ control over their data. The transparency obligation is further reinforced through Article 12 of the Directive, which imposes an obligation on the data controller to provide data subjects information regarding the categories of data under processing, the purpose of the processing, potential recipients and the retention period of the data. However, when data are combined across multiple services based on separate purposes, the word ‘specific’ starts to lose meaning and transparency becomes an inconvenience. This is because the outcome is either information overload or the deliberate use of vague terminology (‘confusology’).¹²⁷

In its opinion on purpose specification, the Article 29 Working Party made it clear that the use of phrases such as “improving users’ experience”, “marketing purposes”, “IT-security purposes” or “future research” will . . . not meet the criteria of being “specific”.¹²⁸ Nevertheless, this is exactly what happened when Google consolidated its services under one account and one privacy policy. In imposing fines on Google, the Spanish DPA indicated that despite the company processing nearly 100 services and products, it failed to provide adequate information about what data it collects and what data are used for which purpose.¹²⁹ Where information is provided, the authority

remarked that Google employed highly vague terminology, such as ‘improving users’ experience’.¹³⁰ The end result of such initiatives is a policy that is ‘indeterminate and unclear’ and of little help to users.¹³¹

In the absence of a sufficiently clear and explicit purpose for processing—in other words, give a lack of transparency—users cannot exercise effective control (eg give meaningful consent) over their data. In a similar investigation conducted by the French DPA, it was indicated that Google’s consolidated privacy policy was considered too vague to form a basis for meaningful consent.¹³² This, in turn, would distort the legal certainty and predictability envisioned by the purpose specification. As noted by the German Constitutional Court,¹³³ in a world where ‘individuals can no longer ascertain who knows what about them’ a lot remains to be desired in the realization of the right to information self-determination.

¹²⁵ WP203 (n 4) 18.

¹²⁶ *ibid* 11.

¹²⁷ This is a practice of firms ‘purposefully introducing uncertainty and confusion into consumer transactions’; see Ryan Calo, ‘Privacy and Markets: A Love Story’, (2015) 91 Notre Dame Law Review 673.

¹²⁸ WP203 (n 4) 13.

¹²⁹ AEPD (n 14).

¹³⁰ *ibid*.

¹³¹ *ibid*.

¹³² Linklaters, ‘EU-Coordinated Enforcement Action over Google’s New Privacy Policy’ (January 2014)

<[http://www.linklaters.com/Insights/Publication1403Newsletter/TMT-News-31-January-2014/](http://www.linklaters.com/Insights/Publication1403Newsletter/TMT-News-31-January-2014/Pages/EU-Coordinated-enforcement-over-Google-s-new-privacy-policy.aspx)

[Pages/EU-Coordinated-enforcement-over-Google-s-new-privacy-policy.aspx](http://www.linklaters.com/Insights/Publication1403Newsletter/TMT-News-31-January-2014/Pages/EU-Coordinated-enforcement-over-Google-s-new-privacy-policy.aspx)>

¹³³ German Constitutional Court on the Census Act 15 December 1983 (BVERFGE 65, 1), cited in Albrecht Philipp, *Hands off Our Data* (Knaur Taschenbuch 2015)

162 _ Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

Furthermore, having a specific and explicit purpose provides evidence of the original purpose, and is thus a necessary prerequisite for ensuring accountability.¹³⁴

This is because the existence of a specific and explicit purpose is essential for the regulators to assess compliance with the relevant principles, such as the legality of the processing, whether the data collected are the minimum necessary, relevant and adequate for the specific purpose, whether the duration of the processing is necessary for the purpose and whether the security measures are appropriate. Similarly, having a specific and explicit purpose allows data subjects to exercise their rights towards the controller.¹³⁵ However, if the purpose of the processing is not communicated to the relevant stakeholders visibly and specifically, it makes it challenging for both regulators and individuals to make the entities accountable for non-compliance. Additionally, if an entity consolidates nearly 100 services into a single account, combining data across these services, any data collection or processing will be relevant either for this or that service within such a large pool of purposes.

Two examples can be illustrative here regarding the accountability problems of aggregation.

Under the EU rules, the processing of sensitive personal data is subject to stricter requirements than non-sensitive personal data. For example, if consent is the basis for processing, for non-sensitive data the consent must be ‘unambiguous’ pursuant to Article 7(a), as opposed to ‘explicit’ consent if it involves the processing of sensitive personal data as per Article 8(2(a)). The latter is a higher standard than unambiguous consent. The stricter requirements are not limited to consent as such, but the ‘appropriateness’ of the security measures pursuant to Article 17 need to be assessed in light of the sensitivity of the data. However, at the centre of the concept of emergent property is that when combined, the data tell a story that one would not be able to glean from the individual components. Accordingly, the combination of non-sensitive personal data could give rise to the emergent property of sensitive personal data.¹³⁶ Such data can only be processed in line with the stricter legal requirements.

This notwithstanding, given that the sensitivity emerges at a later point than the initial collection or exists independent of the individual processing activities based on a specific purpose, controllers might continue to treat such data as non-sensitive and without complying with the stricter requirements. This is not a mere theoretical

claim. Despite its apparent disclaimer for not using sensitive data to serve ads, Google was found to be in breach of the Canadian Health Privacy Law and fined for targeting people with healthcare ads on the basis of their search topics. Like in the EU law, the Canadian law makes a distinction between 'express consent', which is required when processing health (sensitive) data, as opposed to 'implied consent' in 134 WP203 (n 4) 15.

135 *ibid.*

136 Discussing how a combination of innocuous data could reveal intimate details and raise privacy concerns; see Internet Society, 'The Internet of Things: An Overview – Understanding the Issues and Challenges of a More Connected World' (15 October 2015) 26 <<http://www.internetsociety.org/doc/iot-overview>>

Emergence in data privacy: a holistic approach _ 163

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

the case of non-sensitive personal data.¹³⁷ The case demonstrates that Google continued to process sensitive personal data, which emerged from a combination of search queries, without complying with the stricter requirements. Although this case seems to show that the current rules are able to capture the data aggregation practices, its scope remains limited to the individual processing activity for search purposes.

The fact that the aggregation was limited to a single processing activity (ie search) makes it easier to detect non-compliance. In other words, Google's breach was essentially related to the legal basis for the specific processing activity rather than a breach arising from processing activities for multiple purposes. The individualistic nature of the rules means that it becomes more problematic when the sensitivity exists independent of the individual processing—that is, when the sensitivity is a result of the combination of data from different processing operations. This is mainly due to the absence of a framework for looking at the totality of the processing activities and data aggregation practices.

The second accountability-related problem arises where the rights and protections of individuals can vary depending on the legal basis used for collection. An example can be found in the new right of data portability introduced under the newly enacted Regulation. One aspect of the data portability right is to enable the data subject to transmit personal data and other information provided by the data subject from one automated processing system to another.¹³⁸ An essential trigger of this right is that the legal basis for processing should be based on either 'consent' (Article 6(1(a)) or 'contract' (Article 6(1(b))—meaning that processing based on the legitimate interests of the controller or vital interest of the data subject does not give rise to such an obligation. When data initially collected with a diversity of legal bases are aggregated, this obscures the rights of the data subject that are specifically linked to a particular legal basis. For example, in its reply to the French DPA's questionnaire, Google identified the use of one or more of Articles 7(a) (consent), 7(b) (contract) and 7(f) (legitimate interest of the controller) in processing personal data. While the data collected under the first two legal grounds are subject to the data portability rights, the third is not. When data collected through these different legitimate grounds are combined under one account, it might be challenging to distinguish which data are collected on the basis of which legitimate ground and whether those data are subject to the data portability rights or any other right.

Compatibility

The notion of compatible use prohibits the further processing of personal data for purposes that are incompatible with the original purpose. This test prevents the use of data in a manner that exceeds the reasonable expectation of privacy and thereby ensures the foreseeability of the processing.¹³⁹ In this sense, the purpose limitation principle reinforces the principle of fairness.¹⁴⁰ However, the practices of combining

137 Office of the Privacy Commissioner of Canada, 'Report of Findings: Use of Sensitive Health Information for Targeting of Google Ads Raises Privacy Concerns' (14 January 2014) <<https://www.priv.gc.ca/en/opc-actions-and-decisions/investigations/>>

138 GDPR, Art 20(1).

139 Bygrave (n 121) 153.

140 See Rec 38 of the Directive.

164 _ Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user

data across services make it challenging to assess the (in)compatibility of certain uses of data. This is related to the above-described challenge—that is, the fact that the notion of purpose specification (specific, explicitly and legitimate purpose) becomes obscured means that it is either difficult to assess whether certain reuses of data are incompatible with initial purpose or every use of data becomes compatible due to the large pools of purposes and their combinations.

Google-like consolidation initiatives seem to be a strategic evasion of data privacy principles. According to Google, the main goal of the consolidation into one account is to use data collected in one service for providing another service. The difference between the pre- and post-2012 Google policies is all but clearly spelled out by Google itself. According to Google, before 2012, independent services and policies meant that we couldn't combine data from YouTube and search history with other Google products and services to make them better. So if a user who likes to cook searches for recipes on Google, we are not able to recommend cooking videos when that user visits YouTube. . . . We wanted to change that so we can create a simpler, more intuitive Google experience – to share more of each user's information with that user as they use various Google services.¹⁴¹

This sounds more like an open declaration of the breach of the purpose limitation principle. This is surely equivalent to saying that before the 2012 change of policy, Google was actually complying with the purpose limitation principle, but not any longer. The most interesting part of the claim is that the inability to combine data from different services, according to Google, is primarily because its policy was not designed to do so, not because the purpose limitation principle dictates such conduct. Indeed, the further processing of data for another purpose is not by itself a violation of the compatibility aspect, and thus there is the issue of whether the use of search data is or is not compatible with the purpose of YouTube search. Given that the objective of both services is search, there is room for the argument that search data are used to suggest videos on YouTube. However, as noted by the Spanish DPA, the combination and reuse of data collected from one service to another are not limited to Search and YouTube.¹⁴²

In assessing whether the further processing of data is compatible with the original purpose, consideration must be given to different factors, particularly the reasonable expectation to privacy of the data subjects.¹⁴³ However, the idea of having a consolidated privacy framework for 100 different services is in tension with the reasonable expectations of data subjects associated with each processing. Given the different nature of the services, individuals have different expectations of security and privacy, which are often translated into the measures they are willing to take in order to protect their interests. An individual's privacy expectations might not be the same when he/she creates a YouTube account over an e-mail account. Arguably, the individual will have higher expectations of privacy for the e-mail services. This expectation is

¹⁴¹ Google Official Blog (n 17).

¹⁴² AEPD (n 14) 2.

¹⁴³ WP203 (n 4) 16.

Emergence in data privacy: a holistic approach _ 165

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

likely to have effects on the time and effort the individual is willing to invest in protecting his/her privacy and security, including the time and effort related to using strong passwords and selecting privacy preferences.¹⁴⁴ A combination of such accounts associated with different expectations of privacy and security into a single account confuses those different expectations and might result in the user settling for the lower standard of protection.¹⁴⁵

Furthermore, the combination of data from across services or products opens the door for using the data for 'multiple purposes that are not clearly determined', thus violating the restriction on incompatible use. Particularly, the Spanish DPA emphasized the following:

This combination of data across services that allows Google to enrich the personal information it stores, exceeds the reasonable expectations of the average

user, who is not aware of the mass and transversal nature of the processing of their data. Acting in this way Google uses a sophisticated technology that exceeds the capacity of the majority of users to make conscious decisions about the use of their personal information so that, in practice, they lose control over it.¹⁴⁶

The decisions from the Spanish and the French DPAs might seem to suggest that the current rules are able to capture the data aggregation practices, which, in turn, seems to contradict the claim in this article that the individualistic approach is inadequate in light of the aggregation practices. In the above decision, the Spanish DPA was not referring to a breach with respect to a specific processing operation but to how the 'combination of data across services . . . exceeds the reasonable expectations of the average user' and how 'Google uses a sophisticated technology that exceeds the capacity of the majority of users to make conscious decisions'. In this sense, the breach was attributed to the 'totality of the processing operations', rather than to a 'single processing' activity. The reference to 'reasonable expectations' of users implies that the Spanish DPA relied on the fairness principle to condemn Google's conduct. All in all, the decision shows the flexibility of some of the principles, such as the fairness principle, with regard to looking beyond the privacy interests of individuals in the individual processing activities. However, it is questionable that the DPAs intended to depart from the individualistic approach. Perhaps the fact that Google made the bold move to consolidate its different processing activities into seemingly one processing activity under the umbrella of 'Google services' made it easier to apply the fairness principle under these circumstances. If this is true, a more systematic aggregation of data across several processing operations (and without

144 During the 2012 consolidation initiative, Google was criticized for its failure to distinguish 'the innocuous content of search queries . . . from more intrusive types of data such as credit card details or data on telephone communications'; see Linklaters (n 132).

145 See Scott Cleland, 'The Unique Google Privacy Problem: Why Google's Forced Integrated Privacy Policy is So Problematic?' www.Googleopoly.net (25 October 2012) commenting that 'Google's forced integrated privacy policy de-sovereignizes privacy policymaking and pushes the world towards lowest common denominator privacy standards'.

146 AEPD (n 14).

166 _Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

consolidating them under one account) might go undetected. Even if the intention was to depart from the individualistic approach, as noted above, these kinds of encouraging moves towards a holistic approach need to be strengthened by relevant tools and a theoretical framework for their consistent application. The lack of a clear legal and theoretical framework for dealing with such aggregation practices is partly to blame for the uncertainty and uncanny situation that followed Google's announcement to consolidate its processing operations.

It is to be recalled that after Google's consolidation announcement, the Article 29 Working Party asked that the company postpone the consolidation until it could evaluate the implications, which Google ignored.¹⁴⁷ This was followed by the first round of questions intended for Google to clarify its data handling practices, on which Google provided clarification, which was found by the Working Party to be unsatisfactory.¹⁴⁸ The second round of questions was issued, which was again met by unsatisfactory and repetitive answers from Google.¹⁴⁹ Even more, in a move exposing the inadequacy of the current rules in dealing with aggregation across different processing activities, Google questioned the legal basis for such a review by the Working Party and its ultimate aim.¹⁵⁰ In October 2012, the Working Party found Google to be in breach of EU data privacy rules and issued recommendations requesting the implementation of changes to its privacy policies and practices within four months.¹⁵¹ Yet, Google stood firmly and in April 2013, the Working Party announced that Google had not implemented the requested changes. Eventually, this led to fines by national DPAs, including the Spanish and French DPAs.¹⁵²

Therefore, case closed, as one would say. Google was penalized for its breach and similar practices would be handled in the same manner. However, there was one problem. The future compliance dimension was overlooked. This is because Googlelike

aggregations have a perpetual effect on compliance with the data privacy rules in the sense that they distort the very foundation upon which the transparency and accountability mechanisms are built. As noted in section 'Individualistic approach to data privacy', the enforcement of the current rules heavily relies on the ability to distinguish among different processing activities, relate every piece of personal data to a particular processing and then assess the compliance of each processing activity in light of the data privacy principles. In contrast, Google-like aggregation practices blur the lines between different processing activities and make it difficult to attribute every piece of data to a particular processing. Such practices obscure the transparency in terms of what data have been collected, for which purpose(s) and whether the data are necessary for that (those) particular purpose(s). Similarly, assessing whether a certain reuse of data is or is not compatible with original purposes becomes very challenging, which then makes it almost impossible to hold entities

147 Initiative for Competitive Online Market Place (ICOMP), 'The Google Privacy Investigation in Europe: Two Years On' <<http://i-comp.org/wp-content/uploads/2014/10/The-Google-Privacy-Investigation-in-Europe-Two-Years-On4.pdf>>

148 *ibid.*

149 *ibid.*

150 See Google France's Updated Letter Regarding CNIL'S Questioners (20 April 2012) 4

151 ICOMP (n 147) 1.

152 *ibid.* 2.

Emergence in data privacy: a holistic approach _ 167

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

accountable for non-compliance with the data privacy rules. The losses of transparency and accountability, in turn, make it difficult for individuals to understand and enforce their rights.

Thus, unless the measures taken prevent such aggregation from happening or restore the independent processing activities, it is not clear how the individualistic approach continues to regulate such companies, when there is no longer a clear line between the different boxes based on specific and legitimate purposes and the data that belong in those specific boxes. In fact, in light of the above decisions of the Working Party, the Spanish and French DPAs, Google has been in continuous breach of the data privacy rules starting from the date that it consolidated its services, yet, nothing is being done about it. Furthermore, in light of the business value that such aggregation practices add and the diminished transparency and accountability for companies going forward, paying fines for data aggregation practices could become a worthwhile investment. This implies that in the absence of a sufficiently deterring and proportionate measure, such practices could become deliberate strategies for circumventing the application of the data privacy rules. That is where the special responsibility regime discussed in section 'Towards a holistic approach: enhanced responsibility?' might be helpful.

Loss of practical obscurity

In its broadest sense, practical obscurity represents the cost and practical difficulties one encounters in obtaining and compiling information on the private lives of individuals or, more generally, in intruding on a person's privacy. Practical obscurity has served privacy a great deal owing to the costs and difficulties associated with following and recording every footstep an individual takes. However, with the ubiquity of smartphones that are connected to the Internet 24/7 and the sinking cost of storage, companies are able to track and record every digital footprint of an individual through his/her mobile and browsing habits. The more data that are collected, recorded and combined, the more practical obscurity becomes almost a lost cause.

Google-like consolidation and aggregation initiatives are the main culprits regarding the continuous weakening of the importance of practical obscurity. In an era where the capacities of hackers and thieves are equally boosted by computing power, security breaches are not uncommon, even in industries that heavily invest in information security.¹⁵³ In light of such a reality, the consolidation of data across services in general and the single account for many data sets collected for different services do not fit well with information security best practices of 'isolation' and "compartmentalizing information" in order to prevent catastrophic breaches".¹⁵⁴

This means that Google's one account principle creates a single point of failure and vulnerability. A hacker who manages to crack into a YouTube account can easily obtain access to, at the very least, e-mail communications, social network messages and posts, search history and calendar. In a sense, the consolidation reduces the cost of hacking into individual accounts to obtain the same detailed profile and aggravates the impact that a single security breach can have on the rights of individuals.

153 Tene (n 93) 1456.

154 Cleland (n 145).

168 _ Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

Similarly, maintaining anonymity and pseudonymity becomes extremely challenging.

In its 2012 policy, Google declares that 'we may replace past names associated with your Google Account so that you are represented consistently across all our services. If other users already have your email, or other information that identifies you, we may show them your publicly visible Google Profile information, such as your name and photo'.¹⁵⁵ This means that individuals who deliberately use different names and credentials for different accounts can no longer remain anonymous and are forced to be identified despite identification not being necessary to access the service. In this sense, the consolidation represents a regression in the level of data security and further weakens practical obscurity.

This problem is exacerbated when government security agencies access the databases of several organizations all at a certain time ie upstream access. This is partly witnessed in PRISM, the programme that allows US government agencies to access user data from companies such as Google, Facebook, Microsoft and Yahoo. Having reviewed the files obtained from Edward Snowden, Glenn Greenwald reported that the programme grants the NSA direct access to the servers of the participating providers 'without having to request them from the service providers and without having to obtain individual court orders'.¹⁵⁶ According to the Washington Post, 'government employees cleared for PRISM access may "task" the system and receive results from an Internet company without further interaction with the company's staff'.¹⁵⁷ By accessing different data sets at a time, such agencies are able to generate a more detailed profile than what companies such as Google are able to generate from combining data across their wide array of services and products. If, as shown above, the profile from private entities endangers the data privacy rights of individuals through overexpose, access by government authorities to the databases of several such companies that possess detailed profiles of individuals can only exacerbate the problem. More importantly, the capacity to access such profiles opens up the possibility for the government to circumvent constitutional safeguards aimed at protecting privacy.

In a recent article titled 'The Transparent Citizen',¹⁵⁸ Reidenberg cites one example from the USA where the federal government sought information about pornography and issued subpoenas to the five largest search engines ordering them to log files for all user search requests during a specific period.¹⁵⁹ In the absence of such platforms, Reidenberg argues, the alternative would have been to wiretap a large part of the net traffic. For one, this would be a very costly exercise for the government to

155 Google Official Blog (n 17).

156 Glenn Greenwald, 'NSA Prism Program Taps in to User Data of Apple, Google and Others' The Guardian (7 June 2013) <<https://www.theguardian.com/world/2013/jun/06/us-tech-giants-nsa-data>> accessed 30 Sep 2016.

157 Robert O'Harrow Jr, Ellen Nakashima and Barton Gellman, 'U.S., Company Officials: Internet Surveillance Does Not Indiscriminately Mine Data' The Washington Post (8 June 2013) <https://www.washingtonpost.com/world/national-security/us-company-officials-internet-surveillance-does-not-indiscriminately-mine-data/2013/06/08/5b3bb234-d07d-11e2-9f1a-1a7cdee20287_story.html?utm_term=.0f8808499726> accessed 30 Dec 2016.

158 Joel Reidenberg, 'The Transparent Citizen' (2015) 47 Loyola University Chicago Law Journal 101.

159 Most search engines, except for Google, did not even attempt to challenge such large-scale profiling of users; *ibid* 117.

Emergence in data privacy: a holistic approach _ 169

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

conduct.¹⁶⁰ Even more so, this 'is something the government chose to avoid' because

'[i]t would have faced strict legal constraints, namely the need for search warrants for each of the individual account holders.'¹⁶¹ In this sense, the circumvention of the constitutional safeguards—ie the need to obtain a search warrant—is facilitated by the loss of practical obscurity. The loss of practical obscurity, in turn, is facilitated by the accumulation of massive data about individuals within private entities and, more importantly, by their practices of aggregation of data about individuals. In addition, governments can turn their eyes to data aggregating companies, such as ChoicePoint, to connect the missing links.¹⁶² Thus, the accumulation of massive data about individuals and aggregation practices by private entities not only make it cheaper or even cost-free for governments to engage in what could have been a costly conduct but also provide a way of circumventing the safeguards.¹⁶³ The above example shows that having different accounts with different service providers arguably involves some element of practical obscurity because the government had to issue a subpoena to each service provider to access accounts. In a similar fashion, even within the same service provider, one can argue that operating different accounts has some privacy protective feature in that governments, at least in most democratic states, would have to obtain a search warrant for accessing information within each account or at least justify the scope of their request to access all accounts. This means, for example, that before Google's 2012 consolidation, a law enforcement agency that wanted to access someone's Gmail, Search, Google+ and YouTube accounts had to undertake the task of justifying (*ex factum* or *post factum*) whether accessing all four accounts was necessary for the purpose of the investigation. This creates a checkpoint for the judiciary to assess the proportionality of the interference in light of other interests. However, a subpoena after the 2012 consolidation could give the law enforcement full access to someone's Google account—virtually every service the individual uses with Google, as opposed to any specific account. Individuals are even more exposed when governments are able to access a number of similar accounts from different service providers at the same time, as in the case of PRISM.

None of this is to deny that access to such aggregated data increases governments' efficiency in their investigative actions. The argument is not that governments should not request such access from different providers. Rather, given the existence of privacy interests beyond the individual components, any such access should be subject to heightened regulatory or judicial oversight. This means that as the number of processing activities to be aggregated increases so should the safeguards against government interference.¹⁶⁴ Accessing one's Gmail is one thing, but accessing a Google

¹⁶⁰ *ibid.*

¹⁶¹ *ibid.*

¹⁶² Democracy Now, 'Data and Goliath: Bruce Schneier on the Hidden Battles to Collect Your Data and Control Your World' (13 March 2015) <http://www.democracynow.org/2015/3/13/data_and_goliath_bruce_schneier_on>

¹⁶³ *ibid.* (noting that '[a]lthough the federal Privacy Act of 1974 (5 U.S.C. sec. 552a) places some restrictions on uses of personally identifiable information by federal agencies, access to the extensive services of these information brokers appears to be a way around the law').

¹⁶⁴ See Slobogin (n 53); having noted that the duration of surveillance will generate more private information, the author suggests scaleable safeguards against such interference.

¹⁷⁰ *Emergence in data privacy: a holistic approach*

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

account is another. Similarly, accessing one's Google account is frightening, but accessing one's Google, Facebook and Yahoo accounts is completely different. In the face of the increased desire by entities to collect, combine and analyse as much personal data from various sources as possible, ensuring effective data privacy for individuals becomes extremely difficult. However, doing the same against a system that involves the government (NSA), Google, Facebook and Microsoft is even 'more daunting'.¹⁶⁵ Such practices conflate the boundaries between state and economy and thereby distort the checks and balances mechanisms designed to ensure the state's faithfulness to the rule of law.¹⁶⁶ As the lines between governance, surveillance, commerce and private life disappear, the constitutional safeguards for protecting privacy become easy to circumvent.¹⁶⁷ These practices undermine what Bygrave identifies as

Group 2 data privacy interests and more particularly the interests of rule of law and democracy.¹⁶⁸ Further, the loss of practical obscurity is closely related to the interest of non-transparency, which includes the individual's interest 'in being able to act without being identified'.¹⁶⁹

In recognizing the importance of practical obscurity for privacy, the US Supreme Court made a distinction between information that is distributed across different sources and information centralized in one location.¹⁷⁰ According to the Court, '[p]lainly there is a vast difference between the public records that might be found after a diligent search of courthouse files, county archives, and local police stations throughout the country and a computerized summary located in a single clearinghouse of information.'¹⁷¹ The implication of the decision is that the latter represents more risks than the former and thereby warrants more oversight. In a recent decision applying Article 8 of the ECHR, the ECtHR followed a similar approach in *Szabó* and *Vissy v Hungary*. The Court indicated that the need to enhance the 'necessary in a democratic society' test under Article 8(2) to 'strictly necessary' in *Klass* and others¹⁷² was necessitated by 'the particular character of the interference in question and the potential of cutting-edge surveillance technologies to invade citizens' privacy'.¹⁷³ Thus, the Court noted that '[g]iven the technological advances since the *Klass* and *Others* case, the potential interferences with email, mobile phone and Internet services as well as those of mass surveillance attract the Convention protection of private life even more acutely'.¹⁷⁴ The ECtHR further noted that a government's accumulation of 'a detailed profile . . . of the most intimate aspects of citizens' lives may result in particularly invasive interferences with private life'.¹⁷⁵ Having

¹⁶⁵ Harcourt (n 95) 188.

¹⁶⁶ Reidenberg notes that '[t]he transparency of personal information presents a basic challenge to the rule of law both in terms of the state's faithfulness to its ideals and in terms of individuals' respect for its principles'; Reidenberg (n 158) 113.

¹⁶⁷ Harcourt (n 95) 188.

¹⁶⁸ Bygrave (n 5) 150.

¹⁶⁹ *ibid.*

¹⁷⁰ *United States Department of Justice v Reporters Committee* (1989) 489 Supreme Court 764.

¹⁷¹ *ibid.*

¹⁷² *Klass and Others v Germany* App no 5029/71 (6 September 1978).

¹⁷³ *ibid.*, para 73 (emphasis added).

¹⁷⁴ *Szabó and Vissy v Hungary* App no 37138/14 (12 January 2016) para 54 (emphasis added).

¹⁷⁵ *ibid.*, para 70 (emphasis added).

Emergence in data privacy: a holistic approach _ 171

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

noted the sophistication of the surveillance technology, the Court emphasized that '[t]he guarantees required by the extant Convention case-law on interceptions need to be enhanced so as to address the issue of such surveillance practices'.¹⁷⁶ The existence of voluminous data from a wide range of sources, practices of aggregation across these sources and a centralized access point for such data are among the main factors that necessitated the stricter application of the rules, particularly the 'necessity' test. The decision recognizes not only the dangers of the limitless collection of data about an individual, but also the need for a more enhanced and stronger protection than in other circumstances.

Thus, constitutional and legislative measures with the aim of protecting privacy rights should be used for carefully examining the loss of practical obscurity and the overexposure of individuals and for devising appropriate mechanisms for mitigating the dangers through enhanced mechanisms. The argument in this article is that the enhanced mechanisms should not be limited to government data processing practices and should extend to private entities. The subject of the next section is such—an enhanced framework.

TOWARDS A HOLISTIC APPROACH: ENHANCED RESPONSIBILITY?

The main argument of this section is that in light of the interests identified above, the individualistic approach is inadequate and needs to be complemented by a holistic approach with enhanced responsibility for certain actors. This is akin to the regulation under competition law, where companies with a dominant position have a special responsibility and are subject to closer scrutiny and oversight than others.¹⁷⁷

This means that certain restrictions apply only to entities with a dominant position. I argue that data privacy will benefit from a similar framework of scalable regulation that looks at imposing enhanced responsibilities based on the holistic view of the processing activity and data aggregation practices. This approach allows for looking beyond the individual processing activities and considering the wholeness and sum of the processing activities and their associated dangers. Thus, based on the holistic view of the processing activity, more enhanced responsibilities, *ex ante* and *ex post*, might be imposed on certain entities, thereby providing safeguards against the dangers of overexposure and the losses of transparency, accountability and practical obscurity.¹⁷⁸

¹⁷⁶ *ibid* (emphasis added).

¹⁷⁷ Communication from the Commission, 'Guidance on the Commission's Enforcement Priorities in Applying Article 82 (now Art 102) of the EC Treaty to Abusive Exclusionary Conduct by Dominant Undertakings', 2009/C 45/02, para 1 (defining dominance as 'a position of economic strength enjoyed by an undertaking, which enables it to prevent effective competition being maintained on a relevant market, by affording it the power to behave to an appreciable extent independently of its competitors, its customers and ultimately of consumers')

¹⁷⁸ A detailed discussion of concrete policy proposals, in terms what the enhanced responsibility regime is, is beyond the scope of this article. Some of the measures that could be considered might include the obligation for some entities to conduct DPIA when involved in more than a certain number of processing activities and combine data across these processing activities. In the GDPR, these kinds of obligations could easily be incorporated into Art 35(4) where DPAs can establish a list of processing operations which will be subject to DPIA. Further, Rec 92 of the GDPR underlines the need, in certain circumstances, for the DPIA to go beyond a single processing activity. Moreover, as noted in the above section, the ECtHR has upgraded the 'necessity' test to 'strictly necessary' in order to deal with the increased

¹⁷² *Emergence in data privacy: a holistic approach*

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

Under the newly adopted Regulation, the European legislators have already shown an interest in a regulatory approach akin to competition law by introducing significant penalties, which are common under competition law, for the violation of data privacy rules. Interestingly, the fines under the GDPR are scalable and in line with the proposal put forth in this article, as they take into consideration the total worldwide annual turnover.¹⁷⁹ It is also possible that to the extent that competition law applies to control over personal data, for example, in terms of creating a monopoly and concentration, it can help address or mitigate wholeness and sum as a data privacy problem.¹⁸⁰ For example, there are arguments that Google's consolidation initiative and combining of data across services could constitute a form of abusive 'bundling' under competition law.¹⁸¹ If such claims succeed, the typical remedy under competition law is 'unbundling'—i.e. breaking the tie-in and preventing the dominant firm from conditioning the acquisition of one product to the acquisition of another. This would mean that Google cannot offer consolidated services under a single account. This would certainly mitigate some of the privacy concerns with wholeness and sum, as it would reduce the possibility of identifying the user across all services and, to a certain extent, the user's overexposure.

Similarly, the German Competition Law Authority (Bundeskartellamt) recently launched an investigation into Facebook for possible abuse of its dominant position in the market for social networks by imposing unfair terms and conditions on users.¹⁸² Questioning the admissibility of the terms and conditions in light of 'applicable national data protection law', the authority underlined that '[i]f there is a connection between such an infringement and market dominance, this could also constitute an abusive practice under competition law.'¹⁸³ Again, if similar claims

risks of the technological advancement and has further noted that such strictness is needed more 'acutely' in the current technological reality. This approach of 'strictly necessary' could be applied under Arts 6 and 7 of the Directive once an entity is subject to the enhanced responsibility regime. In fact, the 'strictly necessary' test is used under Art 5(3) of the ePrivacy Directive, which prohibits the use of cookies without the consent of the individual, unless the cookie is 'strictly necessary in order for the provider of an information society service explicitly requested by the subscriber or user to provide the service'.

According to the Art 29 Working Party, the 'strictly necessary test' has to pass the following two cumulative tests: (i) A cookie is strictly needed to provide a specific functionality to the user (or subscriber), and if cookies are disabled, the functionality will not be available; and (ii) this functionality has been explicitly requested by the user (or subscriber) as part of an information society service.

¹⁷⁹ See GDPR, Art 83.

¹⁸⁰ There is growing interest in using competition law for big data practices. Recently, the EC Commissioner commented that the DG Committee would examine the extent to which the competition

law rules might apply to big data practices. See Natasha Lomas, 'European Antitrust Chief Eyeing Tech Giants' Hold on Data' (TechCrunch, 18 January 2016) <<https://techcrunch.com/2016/01/18/european-antitrust-chief-eyeing-tech-giants-hold-on-data/>>; see also European Data Protection Supervisor (EDPS), 'Privacy and Competitiveness in the Age of Big Data: The Interplay Between Data Protection, Competition Law and Consumer Protection in the Digital Economy (Preliminary Opinion)' (March 2014).

181 Chilling'Competition, 'How to Distinguish between Tying and Refusal to Deal Cases (Hint: It's Not Just Words)' April 2015 <<https://chillingcompetition.com/2015/04/24/how-to-distinguish-between-tyingand-refusal-to-deal-cases-hint-its-not-just-words/>>

182 See Bundeskartellamt Press Release, 'Bundeskartellamt Initiates Proceeding against Facebook on Suspicion of Having Abused Its Market Power by Infringing Data Protection' <http://www.bundeskartellamt.de/SharedDocs/Meldung/EN/Pressemitteilungen/2016/02_03_2016_Facebook.html>

183 *ibid.*

Emergence in data privacy: a holistic approach _ 173

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

succeed, then Google-like initiatives of one account and its combination of data across services would also come under scrutiny for forcing users to identify across services and imposing unfair terms. Indeed, for these arguments to succeed, they must be supported by adequate economic justifications, which is beyond the purview of this article. Rather, the focus of this article is on examining how the layered approach, based on the special responsibility of entities with a dominant position, can be adopted as relevant to data privacy for dealing with the wholeness and sum problem.

The layered regulatory approach is not entirely new under data privacy, although it is based on a slightly different justification and framework. For example, under the EU data privacy rules, Small and Medium Enterprises (SMEs) are exempted from certain obligations.¹⁸⁴ The main rationale behind the exemption is that SMEs will be deterred from entering the market if they are treated equally with established firms. Thus, the exemptions are there to make it easier for the SMEs to compete with bigger players. However, in reality, research has shown that the current data privacy rules impose an undue burden on SMEs. According to one study, firms with more scope and economy of scale find it easier to obtain consent from users than SMEs do.¹⁸⁵ The idea underlying the argument is that when similar standards of consent are applied across the board, the entities engaged in a smaller scale of processing suffer disproportionately.¹⁸⁶ One possible explanation for the disproportionate costs of 'small and young firms' is the lack of scaling framework as the scope of the processing grows.¹⁸⁷ Given that the enhanced responsibility is primarily based on heightened privacy concerns as a result of the scale and diversity of the information processing activity, it caters not only to the data privacy interests of individuals but also the interests of SMEs.

This idea of scalable responsibility was picked up, albeit very briefly, by the European Data Protection Supervisor (EDPS).¹⁸⁸ In discussing the interplay between consumer protection, data privacy and competition law, the EDPS indicated that the responsibilities under the data privacy rules are incremental depending on the volume, complexity and intrusiveness of a company's personal data processing activities.¹⁸⁹ Citing the ECtHR's decision in *M.M. v UK*,¹⁹⁰ the EDPS strengthened this argument and held that safeguards for protecting personal data should be commensurate in light of the sensitivity of the personal data being processed.¹⁹¹ This

184 For example, under the new Regulation, SMEs are exempted from record keeping obligations; see GDPR, Art 30(5).

185 Avi Goldfarb and Catherine Tucker, 'Privacy and Innovation' in Josh Lerner and Scott Stern (eds), *Innovation Policy and the Economy*, Volume 12 (University of Chicago Press 2012) 83.

186 This further strengthens the suggestion above for the 'strictly necessary' test under Art 7 of the Directive for actors that are subject to the enhanced responsibility regime.

187 Goldfarb and Tucker (n 185) 83.

188 EDPS (n 180).

189 *ibid.* 14.

190 24029/07 Judgment of 13 November 2012.

191 EDPS (n 180) 14. This is not entirely new, because the Directive itself requires that relevant safeguards should be 'appropriate', which is assessed in relation to the 'risks inherent in the processing and the nature of the data to be protected'; see Art 17 and Rec 46 of the Directive. However, this is limited to the security obligations and open to interpretation.

174 _ Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>

incremental responsibility is considered to resemble the special responsibility of dominant undertakings under competition law.¹⁹² This notwithstanding, the application of the enhanced responsibility should be narrowly focused on situations where there are data privacy dangers.

Under competition law, the trigger for the special responsibility regime is the existence of market power and whether an undertaking occupies a dominant position, which is often determined by the market share of the undertaking.¹⁹³ Relevant factors in such an assessment are the undertaking's turnover or volume of total sales.

One may well ask, 'What criteria ought to be used in data privacy law to trigger the enhanced responsibility framework?' The discussion on the concept of emergent property can be informative in identifying different factors that can be taken into account in assessing whether enhanced responsibilities ought to be imposed. However, it bears mentioning that the goal is not to suggest a bright-line rule where the enhanced responsibilities are triggered; neither would it be practicable given the sophistication with which modern business is conducted. Thus, the assessment needs to be conducted on a case-by-case basis, taking into account some of the factors identified here.

In section 'The idea of emergent property in data privacy', it was noted that from a data privacy perspective, emergent property could be looked at from the summative aspect and the interrelation or wholeness aspect. The summative aspect focuses on the scope and amount of information collected and processed based on different purposes that provide insight into different aspects of individual life, regardless of the actual aggregation. One factor that is indicative of the scope is the total number of processing activities which an entity undertakes, which can easily be deduced from the services or products offered by the entity that allow the collection of personal data. This resembles similar criteria of competition law that take account of the economies of scope. These are the economic advantages for a business that could lead to a dominant position, because of the diversity of the goods or services offered and the resulting cost advantages due to the scope of operations.

Under competition law, the special responsibilities regime is justified on the grounds that dominance opens up the potential for stifling competition in a market. In a similar fashion, the scale of processing activities has a direct impact on the privacy interests of individuals, in terms of their overexposure and the losses of accountability, transparency and practical obscurity. As noted above, the larger the number of channels through which an entity collects data from individuals, the more the individuals are exposed and the more difficult it becomes to apply the purpose limitation principle, and, overall, transparency and accountability diminish. Moreover, the more data are collected and stored through a wide array of processing activities relating to individuals, the more the individuals are exposed to governments, as this opens up the possibility for governments to access not only more aggregated and comprehensive data about the individuals in the hands of one entity but also in combination with similar comprehensive data across different service providers. To a certain extent, this shows the link between the total number of processing activities and the

¹⁹² *ibid.*

¹⁹³ Communication from the Commission (n 177).

Emergence in data privacy: a holistic approach _ 175

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>

by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

privacy interests in the wholeness and sum, thereby justifying the application of enhanced responsibility.

Again, the idea is not to suggest a specific number that triggers the enhanced responsibilities; instead, the total number of processing activities could form one criterion for a case-by-case assessment. This means that a certain total number of processing activities can be used as a threshold for starting the assessment to determine whether the enhanced responsibilities should be considered. This minimum threshold would serve the same purpose as market share in the competition law assessment. Under competition law, market share provides a useful first step in the

assessment of dominant position, but other factors should be taken into account, including the nature of the market, entry barriers and potential competition. In this regard, the minimum threshold of total processing activities would bring an entity into the spotlight of conducting further assessments to determine whether it should be subject to the enhanced responsibilities. Certain entities would easily fulfil the criteria. For example, as noted in the sections above, the Spanish DPA investigation indicated that Google collects personal data using more than 100 consumer facing products or services in Spain. The numbers of this scale should provide a strong indication that the enhanced responsibilities should be applied. The same approach is followed under competition law, where a market share of over 50 per cent is considered to provide strong evidence of dominance. However, the total number of processing activities should not be the only criteria that must be factored into the assessment.

One reason for this is that focusing on the number might not capture entities that process personal data under the umbrella of one processing activity, but manage to accumulate massive amounts of personal data either through interrogative techniques (continuous nudging of users to reveal more data)¹⁹⁴ or due to the fact that the usage of the services continuously generates information about the individual. A good example in this regard is Facebook. Although the Facebook service could be considered as a service under a specific purpose (ie providing Facebook services), this does not imply that the dangers of the wholeness and sum are minimal. In fact, Facebook is considered to have a more comprehensive profile of its users than Google does.¹⁹⁵ Thus, if the technology progresses enough to enable the measurement of the size or amount of personal information collected, then it can constitute one factor in the assessment of whether the enhanced responsibilities should be considered.

¹⁹⁶ However, in light of the lack of transparency by entities in terms of what information they collect and the current technological development, it might be difficult to impose obligations according to the amount of data an entity is processing. In this sense, the total number of processing activities provides relatively measurable criteria for assessment. Yet, more general criteria could still be used. For example,

¹⁹⁴ Facebook is a good example because, apart from the troves of personal information it collects through registration and browsing activity, Facebook occasionally comes with new boxes asking users for more information, like 'Do you know a person outside of Facebook?' 'Current job?' 'Family relationship?' And not least, 'What is on your mind?'

¹⁹⁵ See discussions in Harcourt (n 95) 454–55.

¹⁹⁶ This criterion is implicit in Schneier's proposal for fiduciary duties for some actors that control a 'large amount of personal information': Schneier (n 97) 205.

176 Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

whether the entity engages in practices of continuous nagging for information could be considered in such an assessment. In addition, whether the use of the services involves the continuous generation of personal data about the users or whether it is just one-off collection should be taken into account. This is because such features are either undesirable because they constitute nagging or are the root cause of the overexposure of the individual or the losses of accountability, transparency and practical obscurity.

Another aspect of the emergent property is the interrelation that focuses on the aggregation of data across multiple processing activities based on separate purposes. As noted in section 'Privacy interests in wholeness and sum', most of the accountability and transparency challenges are associated with the practice of aggregation or combining data across services. This implies that in combination with the sum of processing activities, the aggregation of data across different processing activities could constitute another consideration for the application of enhanced responsibilities. However, although the existence of actual aggregation sends a strong signal for triggering the enhanced responsibility, its absence should not be taken to reach a contrary conclusion. In other words, the assessment should also take into consideration the potential for the aggregation or combination of data from processing activities based on separate purposes.

Besides those factors that are dictated by the discussions on emergent property,

taking account of other factors, such as the history of privacy breaches, could be relevant. The basis for the application of the special responsibilities for dominant undertakings under competition law is that these undertakings can become 'insensitive to the actions and reactions of competitors, customers and, ultimately, consumers'.¹⁹⁷ Similarly, if an entity is engaged in processing personal data for manifold purposes and has a history of repeated privacy breaches, it shows its insensitivity to compliance and the rights of individuals in general. Such practices ought to warrant enhanced responsibilities. These factors are by no means exhaustive, and other relevant factors could be taken into account depending on the context. The main consideration in the application of the enhanced responsibility regime should be the adequate consideration of the privacy interests identified above—that is, overexposure of individuals and the losses of accountability, transparency and practical obscurity. However, none of this should imply that the proposal put forth is the only alternative for addressing the challenges. It should not also imply that the proposal is without challenges. For example, determining when the privacy interests emerge independent of and/or in combination with the individual processing activities might not be an easy exercise. The aim of this article is rather to highlight the instances where the current rules might benefit from a complementary approach based on a holistic view. Moreover, the objective is to open the floor for further research on similar approaches with a holistic view. Such research could look at what specific measures need to be considered within the enhanced responsibility regime and how they can mitigate the dangers.

197 Communication from the Commission (n 177).

Emergence in data privacy: a holistic approach _ 177

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018

CONCLUSION

This article has shown that when an entity processes personal data for multiple purposes, under the current EU data privacy rules, each processing must comply with the data privacy principles separately, in light of the specific purpose and the relevant legal basis. This represents an individualistic view of the processing—that is, the assessment of compliance focuses on the individual processing based on a specific purpose and distinct legal basis, regardless of the total number of processing activities and data aggregation practices across the different processing operations. In this sense, the EU approach is reductionist. This individualistic (reductionist) approach relies on two underlying assumptions: (i) one is able to distinguish among the different processing activities and relate every piece of personal data to a particular processing, and (ii) if each processing is compliant with the data privacy rules, the data privacy rights of individuals are not endangered.

By drawing lessons from the concept of emergent property, the article has challenged these assumptions and outlined the inadequacy of the individualistic approach in light of emerging data processing practices, where entities process personal data for a panoply of purposes, where almost every use of a service generates personal data, and data are combined across these processing activities. First, although the idea of 'one thing at a time' is a useful guiding principle in life, it might not be suitable in a world where the line between the 'one' and 'the other thing' is unclear. In this regard, the individualist approach overlooks the growing tendency to aggregate or combine data across databases, which makes it difficult to distinguish among processing activities based on separate purposes and to associate certain data with a particular processing activity. Moreover, the concept of emergent property is used to show that when entities process personal data for a wide array of purposes and combine data across different processing activities, there are privacy interests independent of and/or in combination with the individual processing activities. These interests include, but are not limited to, overexposure of the individual and the losses of transparency, accountability and practical obscurity. The discussion through the lens of emergent property has allowed for bringing together issues of surveillance (dataveillance), problems associated with the 'notice and consent' based regulation and some rule of law concerns relating to big data.

Finally, comparisons were drawn from the concept of emergent property in suggesting that data privacy rules need to look beyond the individual processing activities based on a specific purpose and consider the totality of the data collected by entering into a wide variety of new product areas, acquisitions or combinations of data from different sources. It was noted that with an eye open to a holistic approach, there are existing tools that give regulatory authorities flexibility to cater to the interests identified in this article. However, the main goal was not to advocate that the particular possibility sketched here is necessarily the best one available but to spur thinking about complementary approaches that better correspond with the concerns about overexposure and the losses of transparency, accountability and practical obscurity. We must seek to address these concerns, whether be through the holistic approach advanced in this article or any other better option.

178 _ Emergence in data privacy: a holistic approach

Downloaded from <https://academic.oup.com/ijlit/article-abstract/25/2/139/2998971>
by University of Oslo Library. Library of Medicine and Health Sciences user
on 21 February 2018