BRUSSELS PRIVACY HUB
WORKING PAPER
VOL. 6 • N° 24 • JULY 2020

BRUSSELS PRIVACY HUB
Vrije Universiteit Brussel

INDIVIDUATION: RE-IMAGINING DATA PRIVACY LAWS TO PROTECT AGAINST DIGITAL HARMS

Anna Johnston, Principal of Salinger Privacy[1]

Most data protection and privacy laws turn on the identifiability of an individual as the threshold criteria for when data subjects will need legal protection. However I argue that privacy harms can also arise from individuation: the ability to disambiguate or 'single out' a person in the crowd, such that they could, at an individual level, be tracked, profiled, targeted, contacted, or subject to a decision or action which impacts upon them - even if that individual's 'identity' is not known (or knowable). I conclude that data protection and privacy laws need a re-think and re-design in order to reflect the reality of the digital environment, and protect people from digital harms.

First, I will show that 'not identifiable' is no longer an effective proxy for 'will suffer no privacy harm'. Second, I will argue that even the GDPR's mention of 'singling out' is not sufficient to encompass harms arising from individuation. Third, I will demonstrate how some post-GDPR laws and statutory instruments have taken a more expansive approach to threshold criteria, to incorporate individuation. Finally, I will outline a six-part approach which could be taken by legislators to ensure that new or reformed laws robustly protect against digital harms, while avoiding some of the pitfalls demonstrated in the drafting of the CCPA.

Key Words: Personal data, personal information, identifiability, individuation, profiling, GDPR, CCPA, data protection, privacy, AdTech

Contents

The Brussels Privacy Hub Working Papers are intended to circulate research in progress for comment and discussion. Available at www.brusselsprivacyhub.org/publications.html
ISSN N° 2565-9979. This version is for academic use only.

Disclaimer

1. Why definitions matter

Whether or not any particular piece of data meets the definition of 'personal data' is a threshold legal issue for the operation of most privacy and data protection laws (collectively, 'data privacy laws' for the purposes of this paper) around the world. The definition of 'personal data' (or its equivalents such as 'personal information') determines the boundaries of what is regulated, and what is protected, by the privacy principles and data subject rights which follow.[2]

Privacy principles, tempered by exceptions for some scenarios, set out obligations on regulated entities for the handling of personal data, and data subject rights create actionable rights for individuals in relation to the personal data held about them. Data that is not 'personal data' is not subject to the same obligations, or the same protections – even if its collection or use is capable of doing harm to an individual.

Under most data privacy laws, if data does not meet the threshold definition of 'personal data', a dataset can be released as open data, sold to other organisations, or used for a new purpose such as predictive analytics or to train a machine learning system, without legal limits or protections in relation to privacy.

Understanding the scope of what is meant by 'personal data' – and ensuring that that definition remains fit for purpose – is therefore a critical endeavour in privacy jurisprudence.

The definition of personal data (and its equivalents) are in need of a radical re-think and re-design, to ensure they can protect against privacy harms.

2. Are data privacy laws fit for purpose?

Data privacy laws, including the European Union's General Data Protection Regulation (GDPR), have not kept up with rapidly evolving technological advances, and their implications for our privacy – our autonomy, our self-determination and our solitude, our freedom of speech and freedom of association, and the freedom to live without discrimination or fear.

The key problem is that almost all data privacy laws only offer legal protection to individuals who are 'identifiable'.

For example, the Australian Privacy Act turns on the definition of 'personal information', which is:

> "information or an opinion about an **identified** individual, **or** an individual who is **reasonably identifiable**:
> (a) whether the information or opinion is true or not; and
> (b) whether the information or opinion is recorded in a material form or not" (emphasis added).[3]

---

1 Anna Johnston, BA, LLB (Hons I), Grad Dip Leg Prac, Grad Cert Mgmt, MPP (Hons). Salinger Privacy is a privacy consulting and training firm based in Sydney, Australia.

With thanks to Graham Greenleaf, Professor of Law & Information Systems, University of NSW, Australia who reviewed and commented on an earlier draft of this paper. Any mistakes are the author's own.

2 I do note some exceptions, such as the European ePrivacy Directive which is not limited in its scope to 'personal data'; and some Asian laws such as Japan's which can apply obligations to de-identified data as well as identifiable data.

3 Section 6 of the Privacy Act 1988 (Cth)

New Zealand,[4] Canada,[5] the United States[6] and South Africa[7] also have privacy laws applying to 'personal information', drafted with a similar focus on the ***identifiability*** of the information.

European privacy law uses the term "personal data". Under the General Data Protection Regulation (GDPR), this means:

"any information relating to an **identified or identifiable** natural person ("data subject"); **an identifiable person is one who can be identified, directly or indirectly**, in particular by reference to an identifier such as a name, an identification number, location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person" (emphasis added).[8]

Other jurisdictions which use the phrase 'personal data' and turn on the notion of identifiability include Singapore,[9] Hong Kong,[10] Brazil,[11] and the Council of Europe's Convention 108+.[12]

The commonality between these different laws, jurisdictions and legal definitions is that if no individual is identifiable from a set of data, then the relevant privacy principles (or other legal obligations, however expressed) simply won't apply. If no individual can be identified from a dataset, then the dataset can be released as open data, sold to other organisations, or used for a new purpose such as data analytics, without breaching privacy law.

Each of these laws rest on an assumption that privacy harms can only befall an individual who can be identified. That assumption is increasingly being challenged by the realities of the digital economy.

The challenges posed to the effective reach of data privacy laws come from many different directions: new technologies, new interpretations arising from case law, the increasing risks of re-identification, exponential growth in computing power, advances in fields like data analytics and cryptography, the phenomenon of data breaches, the influence of global debates, and new directions in statute law internationally.

4 Section 2 of the New Zealand Privacy Act defines personal information as "information about an identifiable individual".
5 Section 2(1) of the Canadian Personal Information Protection and Electronic Documents Act 2000, which regulates the private sector, defines personal information as "information about an identifiable individual". Section 3 of the Canadian Privacy Act 1985, which regulates the federal public sector, defines personal information as "any identifying number, symbol or other particular assigned to the individual".
6 The Children's Online Privacy Protection Act 1998 (USA) defines personal information as "individually identifiable information about an individual collected online"; see s.312.2, Part 312 of Title 16: Commercial Practices in the Electronic Code of Federal Regulations.
7 The definition of personal information in the South African Protection of Personal Information Act 2013 is "information relating to an identifiable, natural, living person".
8 Article 4, General Data Protection Regulation, Regulation 2016/679 of the European Parliament and of the Council
9 Section 2 of the Personal Data Protection Act 2012 of Singapore defines 'personal data' to mean "data, whether true or not, about an individual who can be identified (a) from that data; or (b) from that data and other information to which the organisation has or is likely to have access".
10 Section 2 of the Personal Data (Privacy) Ordinance of Hong Kong defines personal data "any data (a) relating directly or indirectly to a living individual; (b) from which it is practicable for the identity of the individual to be directly or indirectly ascertained…"
11 Article 5 of the General Data Protection Law 2018 of Brazil defines personal data as "information regarding an identified or identifiable natural person".
12 Article 2 of the Council of Europe's Convention 108+ defines personal data as "any information relating to an identified or identifiable individual".

2.1 Challenges posed by the Internet of Things

Associate Professor Mark Andrejevic and Dr Mark Burdon have written about what they call the 'sensor society', in which the always-on interactive device is also a tool for constant, passive data collection, even in our own homes.[13] Every connected device is capable of being a sensor, and monitoring its users. Even the peak industry body for Internet of Things manufacturers has argued that while each data collection might not be considered personal data in isolation, in combination the data can "yield highly personal information such as home occupancy and a wide range of behaviours".[14] This turns data privacy principles such as collection limitation, and limits on secondary use of data, on their head: "the function is the creep".[15]

De-identification as a method of privacy protection is particularly difficult, if not impossible, in datasets featuring sensor data. Whether from a FitBit, an Amazon Echo, an Apple Watch or an internet-connected vehicle, the rich combination of location data and detailed behavioural data means one individual can be distinguished from millions of other individuals.[16] Associate Professor Nadezhda Purtova has argued that "in increasingly 'smart' environments any information is likely to relate to a person in purpose or effect", if even it is not immediately apparent from its content.[17]

2.2 Challenges posed by location data

With the advent of mobile phones, telephony providers began to know where we were. With the shift to smartphones, that knowledge has spread well beyond just our phone providers; multiple smartphone apps use a mixture of GPS, Bluetooth and Wi-Fi signals to pinpoint locations whenever we carry our phones.

A global 'sweep' of more than 1,200 mobile apps by Privacy Commissioners around the world in 2014 found that three-quarters of all the apps examined requested one or more permissions; the most common was location.[18] Disturbingly, 31% of apps requested information not relevant to the app's stated functionality. A prominent example was a torch app which tracked users' precise location, and sold that data to advertisers.[19]

13 Mark Andrejevic and Mark Burdon, "Detection devices: how a 'sensor society' quietly takes over", *The Conversation*, 5 May 2014; available at http://theconversation.com/detection-devices-how-a-sensor-society-quietly-takes-over-26089

14 IoT Alliance Australia, *Submission to the Australian Competition and Consumer Commission's Digital Platforms Inquiry*, 15 February 2019, p.1; available at https://www.accc.gov.au/system/files/Internet%20of%20Things%20Alliance%20Australia%20%28February%202019%29.PDF

15 Mark Andrejevic and Mark Burdon, "Defining the Sensor Society", *Television and New Media*, Vol 16(1), 2015, pp.19-36; available at https://espace.library.uq.edu.au/view/UQ:326402

16 Scott Peppet, "Regulating the Internet of Things: First Steps Toward Managing Discrimination, Privacy, Security, and Consent", *Texas Law Review*, Vol 93, 2014, p.129; available at https://texaslawreview.org/wp-content/uploads/2015/08/Peppet-93-1.pdf

17 Nadezhda Purtova, "The law of everything. Broad conception of personal data and future of EU data protection law", 2018, *Law, Innovation and Technology*, Vol 10(1), pp.40-81; available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3036355.

18 Office of the Privacy Commissioner of Canada, "From APP-laudable to dis-APP-ointing, global mobile app privacy sweep yields mixed results", 9 September 2014; available at https://www.priv.gc.ca/en/blog/20140909/

19 See https://www.ftc.gov/news-events/press-releases/2013/12/android-flashlight-app-developer-settles-ftc-charges-it-deceived

However it is not only apps we install on our mobile phones which can track our location. Bluetooth signals emitted by wearable devices can be collected by third parties; and venues such as shopping centres and airports (or, briefly, rubbish bins in London)[20] use the MAC addresses broadcast by devices to detect how populations are moving within a space, and to identify repeat visitors.[21]

Bluetooth Beacons can also be used to link online advertising to offline transactions. Having purchased MasterCard transaction data in the US to better tie offline purchases with online advertisements,[22] Google offers advertisers the ability to see whether an ad click or video view results in an in-store purchase within 30 days.[23] Connecting to shopping centre Westfield's free wifi involves agreeing to a set of terms and conditions which include linking the mobile device ID with the individual's wifi use.[24]

Location data is highly granular. One study suggested that four points of geolocation data alone can potentially uniquely identify 95% of the population.[25] Mark Pesce, a futurist, inventor and educator, has described the geolocation data collected by and broadcast from our smartphones as "almost as unique as fingerprints".[26]

Data showing where a person has been can reveal not only the obvious, like where they live and work or who they visit, but it may also reveal particularly sensitive information – such as if they have spent time at a church or a needle exchange, a strip club or an abortion clinic. Some app-makers claim they can even tell which floor of a building people are on.[27]

A recent example is the analysis conducted by Singaporean company Near on the movements of workers at an abattoir in Melbourne, which was the centre of an outbreak during the COVID-19 isolation period.[28] Near claimed that it could track this small cohort of workers to specific locations including shops, restaurants and government offices. (Near uses "anonymous mobile location information" collected "by tapping data collected by apps" to provide insight into the precise movements of individuals, in order to offer advertisers "finer slices of audiences to reach highly qualified prospective customers".[29] Near boasts of having "the world's largest data set of people's behavior in the real-world" consisting of 1.6 billion 'users', across 44 countries, processing 5 billion events per day.[30])

20 "U.K. bars trash cans from tracking people with Wi-Fi", CBS News, 12 August 2013; available at https://www.cbsnews.com/news/uk-bars-trash-cans-from-tracking-people-with-wi-fi/

21 Jules Polonetsky and Elizabeth Renieris, Future of Privacy Forum Whitepaper: "*Privacy 2020: 10 Privacy Risks and 10 Privacy Enhancing Technologies to Watch in the Next Decade*", January 2020, p.4; available at https://fpf.org/wp-content/uploads/2020/01/FPF_Privacy2020_WhitePaper.pdf

22 See BBC News Online, "Google and Mastercard in credit card data deal", 31 August 2018; available at https://www.bbc.co.uk/news/technology-45368040

23 See https://support.google.com/google-ads/answer/6190164?hl=en-GB

24 See https://www.westfield.com.au/terms-and-conditions#wi-fi-terms-of-use-and-privacy-terms

25 Yves-Alexandre de Montjoye, César A. Hidalgo, Michel Verleysen & Vincent D. Blondel, 'Unique in the Crowd: The privacy bounds of human mobility', Scientific reports, March 2013, available at: https://www.nature.com/articles/srep01376?ial=1

26 Mark Pesce was keynote speaker at the OAIC Business Breakfast for Privacy Awareness Week in May 2015; this quote is from the author's contemporaneous notes from the event.

27 David Pierce, "Location Is Your Most Critical Data, and Everyone's Watching", *Wired*, 27 April 2015; available at https://www.wired.com/2015/04/location/

28 See https://blog.near.co/news/workers-tracked-20km-from-infected-abattoir/

29 See https://blog.near.co/news/we-know-which-suburb-eats-more-pizza-by-analyzing-data-from-15-million-australians/

30 See https://near.co/data/

This information can then be used to target individuals. For example anti-abortion activists use geo-fencing to target online ads at women as they enter abortion clinics.[31] Near has reported that it could target individuals with messaging about the Australian Government's COVIDSafe app: "We can support app adoption, saying to someone you've been to a postcode or a high-risk area and encourage them to download the app. That's quite easy to do".[32] This is despite the company's claim that its data is "anonymized to protect privacy".

None of these technologies – or their ability to impact on people's private lives or autonomy – depend on the identifiability of the data subject. Nonetheless digital platforms, publishers, advertisers, ad brokers and data brokers claim to work outside the reach of privacy laws because the data in which they trade is 'de-identified' or 'anonymised' or 'non-personal'.[33]

2.3 Challenges posed by data analytics

Further, advances in data analytics, and the predictive capabilities of machine learning and artificial intelligence technologies, are also creating new challenges for the law's ability to draw a bright line between what is 'personal data' and what is not.

Professor Sandra Wachter has critiqued the GDPR's framing of personal data and non-personal data as "outdated (and) ineffective" because privacy and discrimination harms can still occur based on user profiles built on non-identifying data.[34]

Philosopher and mathematician Rainer Mühlhoff has similarly argued that the distinction between identifiable and non-identifiable data is no longer effective because "high-resolution yet anonymous mass data" is being used for predictive algorithmic decision making: "Algorithms that are suitable for the management of whole populations based on behavioral data are not concerned with names and identities. … the societal risk associated with Big Data is not identification or disclosure of personal information, but the ***algorithmic selection of societal groups*** that are treated differently in terms of access to opportunities, resources and information".[35]

The Office of the Victorian Information Commissioner has noted that:

> "the distinction between what is and is not considered to be 'personal' is being challenged by the increasing ability to link and match data to individuals, even where previously thought to be 'de-identified' or non-identifying to begin with.
>
> … a combination of seemingly non-personal information can become personal information when analysed or correlated. As the amount of available data increases, and technologies for processing and combining it improve, it becomes increasingly difficult to assess whether a given piece of data

31 Marc Faletti, "How Geo-Fencing Works… and How It Can Be Abused", *Rewire News*, 25 May 2016; available at https://rewire.news/videos/2016/05/25/geofencing-works-can-abused/

32 See https://blog.near.co/news/workers-tracked-20km-from-infected-abattoir/

33 Dr Katharine Kemp, "Submission in Response to the Australian Competition and Consumer Commission Ad Tech Inquiry Issues Paper", 26 April 2020; available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3587239

34 Professor Wachter, "Data Protection in the Age of Big Data", *Nature Electronics*, Vol 2 (6–7), April 2019; available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3355444

35 Rainer Mühlhoff, "We Need to Think Data Protection Beyond Privacy", *Medium*, 10 April 2020; available at https://medium.com/@rainermuehlhoff/why-we-need-data-protection-beyond-privacy-aba9e9c996ed

is 'identifiable'; considering a piece of data in isolation is not compatible with AI technology, and is no longer a true reflection of whether it can be deemed 'personal information'".[36]

Professor Sandra Wachter and Dr Brent Mittelstadt argue that European data protection law, which is "meant to protect people's privacy, identity, reputation and autonomy", is nonetheless "currently failing" because it does not protect individuals from "inferential analytics", because inferences do not necessarily meet the definition of 'personal data'.[37] (By contrast, post-GDPR laws in other jurisdictions have explicitly included inferred data.)[38]

2.4 Current definitions are no longer fit for purpose

Data privacy laws which utilise a definition of 'personal data' turning on identifiability no longer offer a legal framework suitable for the challenges of the digital age.

Professor Sandra Wachter and Dr Brent Mittelstadt argue that "identifiability as a prerequisite to exercise individual rights creates a gap in the protection afforded to data subjects against inferential analysis". They propose that the potential for privacy harms should be reflected in future jurisprudence, "regardless of whether the affected parties can be identified".[39]

The reason we have data privacy laws is not to protect data; it is to protect people. It is the people who can be found in data, singled out because of data, then tracked, profiled, targeted and even manipulated via data, who matter. It is because of the scope to do harm to people that some practices are deserving of regulation.

Privacy harms exist across a spectrum, and include:

- **tangible** or 'material' harms at one end (such as physical harm or threats of violence, stalking and harassment, identity theft, financial loss and psychological damage),
- **intangible** or 'moral' harms in the middle (such as reputational damage, "creepy inferences", humiliation, embarrassment or anxiety, loss of autonomy, discrimination and social exclusion), and
- **shared** or 'social' harms at the other end (such as the threats to democracy, chilling effect on free speech, loss of trust and social cohesion posed by a 'surveillance society', and by manipulation and amplification of political messaging on social media).[40]

36 Office of the Victorian Information Commissioner, "Artificial intelligence and privacy: Issues Paper", June 2018, p.9; available at https://ovic.vic.gov.au/resource/artificial-intelligence-and-privacy/

37 Professor Wachter and Dr Mittelstadt, *"A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI",* Columbia Business Law Review, 2019(2); available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3248829

38 For example, s.3(28) of India's 2019 *Personal Data Protection Bill* includes "any inference drawn from such data for the purpose of profiling"; available at https://www.prsindia.org/sites/default/files/bill_files/Personal%20Data%20Protection%20Bill%2C%202019.pdf. The California Consumer Privacy Act explicitly includes, within its definition of 'personal information', "purchasing or consuming histories or tendencies", and "inferences drawn from" any of the other types of information enumerated in the definition, "to create a profile about a consumer reflecting the consumer's preferences, characteristics, psychological trends, predispositions, behavior, attitudes, intelligence, abilities, and aptitudes": CCPA section 1798.140(o)(1), parts (D) and (K); the full text of the CCPA, as at 1 January 2020, is available at http://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?lawCode=CIV&division=3.&title=1.81.5.&part=4.&chapter=&article=

39 Professor Wachter and Dr Mittelstadt, *"A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI",* Columbia Business Law Review, 2019(2); available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3248829

40 This spectrum of privacy harms is drawn from speeches by the former UK Information Commissioner Richard Thomas, as well as the Future of Privacy Forum's paper, "Benefit-Risk Analysis for Big Data Projects", September 2014, available at www.futureofprivacy.org

The assumption upon which most data privacy laws rest is that identifiability is the key to harm: for example "the underlying conceptual focus of defining personal information in Australian privacy laws regards the revealment of identity as the social harm to be protected".[41]

In other words, the assumption is that no harm can befall an individual from the handling of their personal data if they cannot be ***identified*** from the data; that information which might otherwise cause embarrassment, humiliation, or physical, psychological or financial risks cannot cause such harms if no-one knows ***who*** the information is about.

However in the 21st century, that assumption is no longer true.

I argue that privacy harms can arise regardless of whether or not the person is identifiable in a concrete or legally verifiable sense. In other words, a perpetrator can hurt someone without ever knowing who they are.

Some jurisdictions such as the GDPR have definitions of 'personal data' (or its equivalent), which clearly anticipate device identifiers, online identifiers and location data being used to 'indirectly identify' individuals. However they still depend on an individual ultimately being findable and identifiable in a legal sense, however many steps are required to achieve that, so long as those steps are not legally prohibited.[42]

By contrast, some newer, post-GDPR statues and legal instruments are dramatically broadening out the notion of 'identifiability' (or even abandoning it altogether) as the threshold element of their definition.

I propose that a re-think and re-design of the scope of data privacy laws is necessary to enable comprehensive legal protections from privacy harms. In particular, I argue that the concept of ***individuation***, as well as identification, must be explicitly incorporated into data privacy laws, in order to enable legal protections against digital harms.

3. Individuation

From the digital breadcrumbs we leave behind in the form of geolocation data shed from our mobile devices, to the patterns of behaviour we exhibit online as we browse, click, comment, shop, share and 'like', we can be tracked. Tracked; then profiled; and finally targeted … all without the party doing the tracking, profiling or targeting needing to know 'who' we are.

By linking a device to behaviour such as searches, queries, posts, browsing sites and purchases, the party doing the tracking can start to profile individuals, drawing inferences about their interests and preferences, behaviour and budget, and divide them into segments accordingly. The individual presumed to be the user of the device can then be targeted to receive a particular ad, offered personalised content or recommendations, sent political messaging, or subjected to an automated decision such as differential pricing.

41 Mark Burdon and Paul Telford, "The Conceptual Basis of Personal Information in Australian Privacy Law", *eLaw Journal: Murdoch University Electronic Journal of Law*, 2010, Vol 17(1), p.27; available at https://eprints.qut.edu.au/37696/

42 See the CJEU's decision about indirect identification via dynamic IP addresses in the 2016 *Breyer* case, which was determined under the 1995 Data Protection Directive: see Case C-582/14, *Patrick Breyer v. Bundesrepublik Deutschland* [2016] ECLI:EU:C:2016:779; available at http://curia.europa.eu/juris/document/document.jsf?docid=184668&doclang=EN

Professor Sandra Wachter and Dr Brent Mittelstadt argue that "companies' widespread implementation of inferential analytics for profiling, nudging, manipulation, or automated decision-making … (can) impact the privacy of individuals"; "inferences drawn from anonymous and non-personal data still pose risks for data subjects", yet are excluded from the scope of European data protection law.[43]

The digital environment has turned on its head the assumption that identifiability – in the sense of knowing a person's 'identity' - is the only vector for privacy harm. As the Office of the Australian Information Commissioner (OAIC) has noted, "harm can be caused by just knowing attributes of an individual, without knowing their identity".[44]

3.1 Harms caused without identifiability

One disturbing recent example is the finding that publicly disclosed de-identified data about public transport cards used in the city of Melbourne, could be used to find patterns showing young children travelling without an accompanying adult. Those children could be targeted by a violent predator as a result, without the perpetrator needing to know anything about the child's identity.[45] Other examples of potential harm arising even without identities being revealed have included the release of data about taxi trips[46] and consumers' fitness routines.[47]

If the objective of data privacy laws is to protect people's privacy, those laws need to grapple with a broader view of the types of practices which can harm privacy – regardless of whether an individual's identity is known or revealed.

This paper uses the word ***individuation*** to refer to the ability to disambiguate or 'single out' a person in the crowd, such that they could, at an individual level, be tracked, profiled, targeted, contacted, or subject to a decision or action which impacts upon them - even if that individual's 'identity' is not known (or knowable).[48]

Individuation is the technique used in online behavioural advertising; advertisers don't need to know who any particular consumer is, but if they know that the user of a particular device has a certain collection of attributes, they can target or address their message to the user of that device accordingly.

The objective of online behavioural advertising is, like any advertising, to predict purchasing interests, and drive purchasing decisions. Online, however, the repercussions are much greater, because of the

43 Professor Wachter and Dr Mittelstadt, *"A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI"*, Columbia Business Law Review, 2019(2); available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3248829

44 Office of the Australian Information Commissioner, *What is personal information?*, May 2017, p.21; available at https://www.oaic.gov.au/agencies-and-organisations/guides/what-is-personal-information

45 Dr Chris Culnane, Associate Professor Benjamin Rubinstein, and Associate Professor Vanessa Teague, "Two data points enough to spot you in open transport records", *Pursuit*, University of Melbourne, 15 August 2019; available at https://pursuit.unimelb.edu.au/articles/two-data-points-enough-to-spot-you-in-open-transport-records

46 Anthony Tockar, "Riding with the Stars: Passenger Privacy in the NYC Taxicab Dataset", Neustar blog, 15 September 2014; available at https://research.neustar.biz/2014/09/15/riding-with-the-stars-passenger-privacy-in-the-nyc-taxicab-dataset/

47 Rosie Spinks, "Using a fitness app taught me the scary truth about why privacy settings are a feminist issue", Quartz, 1 August 2017; available at https://qz.com/1042852/using-a-fitness-app-taught-me-the-scary-truth-about-why-privacy-settings-are-a-feminist-issue/

48 Anna Johnston, "Individuation – Re-thinking the scope of privacy laws", 30 August 2016, Salinger Privacy blog; available at https://www.salingerprivacy.com.au/2016/08/30/individuation/

degree to which advertising – and indeed, the very content users are shown – has become 'personalised'. Personalisation means decisions are made about who sees what, and equally what will be withheld from whom.

By allowing exclusion, digital platforms also allow discrimination. Facebook has been caught allowing advertisers to target – and exclude – people on the basis of their 'racial affinity', amongst other social, demographic, racial and religious characteristics.[49] For example, a landlord with an advertisement for rental housing could prevent people profiled as 'single mothers' from ever seeing their ad; an employer could prevent people identifying as Jewish from seeing a job ad; or a bank could prevent people categorised as 'liking African American content' from seeing an ad for a home loan.[50]

Existing patterns of social exclusion, economic inequality, prejudice and discrimination are further entrenched by micro-targeted advertising, which is hidden from public view and regulatory scrutiny. Preying on vulnerable individuals which could lead to physical, financial or social harm is also a risk of micro-targeting. For example a pharmaceutical company selling addictive opioid-based pain medication used Google's search terms data to target people with chronic pain, promoting ads of escalating intensity across multiple sites, despite laws prohibiting the advertising direct to consumers of prescription medication.[51] It was also revealed in 2017 that Australian Facebook executives were promoting to advertisers their ability to target psychologically vulnerable teenagers.[52]

'Personalisation' can lead to price discrimination, like pricing based on an airline knowing this user has searched for a quote before; or market exclusion, like insurance products only being advertised to users already profiled as 'low risk', based on their online activities.[53] Micro-targeting can also be used to manipulate behaviour, such as voting intentions.[54]

The Facebook / Cornell University research project on emotional contagion, revealed in 2014, offers another fine example of causing privacy harm, without 'personal data' being involved. The project deliberately manipulated the news feeds of almost 700,000 unsuspecting Facebook users, and monitored their reactions, in order to trigger emotional outcomes for people who had no idea they were even part of a 'research' project. The researchers described their project as "a massive (*N* = 689,003) experiment on Facebook", which demonstrated "that emotional states can be transferred to others via emotional contagion, leading people to experience the same emotions without their awareness".[55]

---

49 Julia Angwin, Ariana Tobin and Madeleine Varner, "Facebook (Still) Letting Housing Advertisers Exclude Users by Race", *ProPublica*, 17 November 2017; available at https://www.propublica.org/article/facebook-advertising-discrimination-housing-race-sex-national-origin

50 Alex Hern, "Facebook lets advertisers target users based on sensitive interests", *The Guardian*, 16 May 2018; available at https://amp.theguardian.com/technology/2018/may/16/facebook-lets-advertisers-target-users-based-on-sensitive-interests

51 Alison Branley, "Google search data used by pharma giant to bombard users with ads for addictive opioids", ABC News Online, 13 July 2019; available at https://www.abc.net.au/news/2019-07-13/searches-data-mined-by-pharma-giant-to-promote-new-opioid/11300396

52 Nitasha Tiku, "Get Ready for the Next Big Privacy Backlash Against Facebook", *Wired*, 21 May 2017; available at https://www.wired.com/2017/05/welcome-next-phase-facebook-backlash/

53 Rafi Mohammed, "How Retailers Use Personalized Prices to Test What You're Willing to Pay", *Harvard Business Review*, 20 October 2017; available at https://hbr.org/2017/10/how-retailers-use-personalized-prices-to-test-what-youre-willing-to-pay

54 Luke Dormehl, "Will Your Computer Tell You How to Vote?", *Politico Magazine*, 25 November 2014; available at https://www.politico.com/magazine/story/2014/11/computers-algorithms-tell-you-how-to-vote-113142

55 Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock, "Experimental evidence of massive-scale emotional contagion through social networks", *Proceedings of the National Academy of Sciences of the United States of America*, 17 June 2014; available at https://www.pnas.org/content/111/24/8788.full

The university researchers argued that because they did not know who their research subjects were, there was no identifiable personal data at stake. On that basis, Cornell University's Institutional Review Board had concluded that the academics were "not directly engaged in human research", and therefore no ethical review was required.[56] Clearly users did not see the issue this way, with typical reactions being descriptions such as 'creepy', 'evil, 'terrifying' and 'super disturbing'.[57]

3.2 These harms are privacy harms

The UN's Special Rapporteur on Privacy, Joe Cannataci, has written about privacy as enabling the free, unhindered development of personality.[58] Privacy is related to the right to self-determination, and is a critical element of autonomy.

The activities described above hold the potential to impact on individuals' autonomy, by narrowing or altering their market or life choices. Philosophy professor Michael Lynch has said that "taking you out of the decision-making equation" matters because "autonomy enables us to shape our own decisions and make ones that are in line with our deepest preferences and convictions. Autonomy lies at the heart of our humanity".[59]

A person does not need to be identified in order for their autonomy to be undermined or their dignity to be damaged.

Much effort is expended by advertisers and others wishing to track people's movements and behaviours, whether offline or online, in convincing privacy regulators and consumers that their data is not identifying, and that therefore there is no cause for alarm. Whether in double-blind data matching models or the use of homomorphic encryption to compare data from multiple different sources (the sharing of which would be prohibited if the data was 'identifiable'), the current obsession is how to avoid *identifying* anybody, such that the activity can proceed unregulated by data privacy laws. In fact the real question both companies and governments should be asking is how to avoid *harming* anybody.

Security and privacy academic and writer Bruce Schneier has argued that laws concerned with identifiability as the key element are too limiting in their treatment of potential harm:

> "most of the time, it doesn't matter if identification isn't tied to a real name. What's important is that we can be consistently identified over time. We might be completely anonymous in a system that uses unique cookies to track us as we browse the internet, but the same process of correlation and discrimination still occurs. It's the same with faces; we can be tracked as we move around a store or shopping mall, even if that tracking isn't tied to a specific name."[60]

56 Robinson Meyer, "Everything We Know About Facebook's Secret Mood Manipulation Experiment", *The Atlantic*, 28 June 2014; available at https://www.theatlantic.com/technology/archive/2014/06/everything-we-know-about-facebooks-secret-mood-manipulation-experiment/373648/

57 "Facebook probed on mood experiment", *SBS News*, 2 July 2014; available at https://www.sbs.com.au/news/facebook-probed-on-mood-experiment

58 Joe Cannataci, "Privacy, personality and flows of information – an open invitation", personal blog, 9 June 2016; available at https://www.privacyandpersonality.org/2016/06/privacy-personality-and-flows-of-information-an-open-invitation/

59 Michael Lynch, "Why does our privacy really matter?", *Christian Science Monitor*, 22 April 2016; available at https://www.csmonitor.com/World/Passcode/Security-culture/2016/0422/Why-does-our-privacy-really-matter

60 Bruce Schneier, "We're banning facial recognition. We're missing the point", *New York Times*, 20 January 2020; available at https://www.nytimes.com/2020/01/20/opinion/facial-recognition-ban-privacy.html

The Office of the Privacy Commissioner of Canada (OPCC) has also taken a broad, contextual view of the definition of 'personal information', when considering online behavioural advertising. Although noting that this view about identifiability should be tested on a case-by-case basis, the OPCC's view is that:

> "in the context of (online behavioural advertising), given the fact that the purpose behind collecting information is to create profiles of individuals that in turn permit the serving of targeted ads; given the powerful means available for gathering and analyzing disparate bits of data and the serious possibility of identifying affected individuals; and given the potentially highly personalized nature of the resulting advertising, it is reasonable to take the view that the information at issue in behavioural advertising not only implicates privacy but also should generally be considered 'identifiable' in the circumstances".[61]

The OPCC has further noted that their consultations found general agreement amongst stakeholders that "there were privacy implications from online tracking (even if not all agreed that the data collected from tracking was personal information)".[62] Whether or not identity could be determined from the information gleaned through online tracking was the sticking point, in terms of meeting the legal definition for 'personal information'.

If the end result of an activity is that an individual can be ***individuated*** from a dataset, such that they could, at an individual level, be tracked, profiled, targeted, contacted, or subject to a decision or action which impacts upon them, that is a privacy harm which may need protecting against. (Whether or not any particular conduct should be then prevented by the application of data privacy laws will depend on the context; for example an intervention for fraud prevention or crime detection is a different proposition to online behavioural advertising. It is within the more detailed privacy principles that each data privacy law defines the allowable purposes for the collection, use or disclosure of personal data).

In the digital environment, 'not identified' is no longer an effective proxy for 'will suffer no privacy harm'. I propose that individuation should be anticipated by, and explicitly built into, data privacy laws as well.

4. A new definition is needed

Purtova has argued that "the success of future legal protection against 'information-induced harms'" will depend on "a better understanding of information and its relationship to people".[63] In particular, she argues that the "duality" caused by the threshold definition of 'personal data' – data about identifiable individuals is protected, while the rest is not – "is at odds with the world where any information has the potential to affect people".

While incorporating the concept of individuation into the threshold criteria for the application of data privacy laws might appear novel at first, I argue that it is simply an evolution of the concept of identifiability.

61 Office of the Privacy Commissioner of Canada, "Policy position on online behavioural advertising", December 2015; available at https://www.priv.gc.ca/en/privacy-topics/technology/online-privacy-tracking-cookies/tracking-and-ads/bg_ba_1206/

62 Office of the Privacy Commissioner of Canada, "Policy position on online behavioural advertising", December 2015; available at https://www.priv.gc.ca/en/privacy-topics/technology/online-privacy-tracking-cookies/tracking-and-ads/bg_ba_1206/

63 Nadezhda Purtova, "The law of everything. Broad conception of personal data and future of EU data protection law", 2018, *Law, Innovation and Technology*, Vol 10(1), pp.40-81; available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3036355.

By appreciating that the concept of 'identifiability' is not limited to the ability to know or verify a person's legal ***identity***, but that it can also encompass the act of 'singling out' – i.e. disambiguating one individual from another by way of either recognising or assigning certain characteristics to each person - we are already embracing the concept of individuation.

The concept of individuation is arguably already hinted at in the GDPR, given:

- the definition of 'personal data' includes the concept of both direct and 'indirect' identification[64]
- online identifiers and location data are explicitly mentioned within the definition of 'personal data' as possible identifiers, by reference to which an individual might be "identified, directly or indirectly".
- the concept of online identifiers is explicated further in Recital 30 to include IP address, cookies and RFID tags where used to create profiles of people and identify them;[65] and
- Recital 26 mentions 'singling out' as a means by which someone might become 'identifiable'.[66]

Professors Paterson & McDonagh, referring to Recital 26 in the GDPR, conclude:

> "The express reference to 'singling out' suggests that the processing of data that singles out but does not reveal an individual's identity comes within the scope of European data protection law".[67]

However each of these elements still comes back to the idea of the person ultimately being identifiable in a legal sense. Case law from the Court of Justice of the European Union (CJEU) does not assist in shedding light on whether 'singling out' ***without*** identifiability is within scope of the GDPR. While the ***Breyer*** case about indirect identification is held up as an example of some expansive thinking, it rested on the complainant Breyer being ultimately findable and identifiable in a legal sense; the Court found that he was identifiable via the assistance of a third party, so long as the steps necessary to achieve that identification were not legally prohibited.[68] The Court has not directly addressed an example of an individual whose identity was ***not*** knowable.

As noted above, Professor Sandra Wachter and Dr Brent Mittelstadt argue that European data protection law, which is "meant to protect people's privacy, identity, reputation and autonomy", is nonetheless "currently failing" because it does not protect individuals from "inferential analytics".[69] They have highlighted ways in which the CJEU has been inconsistent in assessing the scope of 'personal data', such that inferences drawn about even identifiable individuals has been excluded from scope.

---

64 Article 4, *General Data Protection Regulation*, Regulation 2016/679 of the European Parliament and of the Council

65 Recital 30 of the GDPR states: "Natural persons may be associated with online identifiers provided by their devices, applications, tools and protocols, such as internet protocol addresses, cookie identifiers or other identifiers such as radio frequency identification tags. This may leave traces which, in particular when combined with unique identifiers and other information received by the servers, may be used to create profiles of the natural persons and identify them."

66 Recital 26 of the GDPR includes: "To determine whether a natural person is identifiable, account should be taken of all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly."

67 Moira Paterson & Maeve McDonagh, "Data Protection in an era of Big data: the Challenges posed by Big Personal Data", *Monash University Law Review*, Vol 44(1), 2018, p.16; available at https://www.monash.edu/__data/assets/pdf_file/0009/1593630/Paterson-and-McDonagh.pdf ; see also Dr Frederik J. Zuiderveen Borgesius, "Singling out people without knowing their names – Behavioural targeting, pseudonymous data, and the new Data Protection Regulation", *Computer Law & Security Review*, Vol 32(2), April 2016, pp.256-271; available at https://www.sciencedirect.com/science/article/pii/S0267364915001788?via%3Dihub

68 See http://curia.europa.eu/juris/document/document.jsf?docid=184668&doclang=EN

69 Professor Wachter and Dr Mittelstadt, *"A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI",* Columbia Business Law Review, 2019(2); available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3248829

Dr Monique Mann and Professor Tobias Matzner have also highlighted the limitations of the GDPR in protecting against emerging forms of discrimination, because of its focus on harms to identifiable individuals.[70]

4.1 Newer thinking in drafting privacy statutes

Dr Katharine Kemp has argued that "identification should not be limited to data which is labelled with a consumer's legal name or contact details, but should extend to data which can be used to single out one consumer as distinct from other consumers".[71]

Indeed, some statutes and other international instruments drafted since the GDPR have shifted towards more explicitly incorporating the concept of individuation, moving beyond just ***identifiability*** as the essential threshold element, to also incorporate notions such as data which can be used to ***reflect***, ***recognise***, ***contact*** or ***locate*** an individual.

Examples are:

- The 2017 Information Security Standard in China, which fleshes out the expectations of how the 2016 Cybersecurity Law applies in practice: in addition to taking a 'capacity to identify' approach to its threshold definition, the Standard also incorporates data which 'can **reflect the activities** of a natural person' (emphasis added). Privacy academic Professor Graham Greenleaf and IP lawyer Scott Livingston describe this second possible element as "a fairly expansive broadening away from information that 'identifies' an individual to any information that may 'reflect' a specific person (without necessarily identifying) them". Greenleaf and Livingston further suggest that the effect is to regulate data "which gives the organisation the capacity to interact with a person on an individuated basis" using data which does not necessarily enable the data subject to be identifiable.[72]
- The 2018 California Consumer Privacy Act (CCPA) expressly includes, within its definition of personal information, data which is "**capable of being associated with**, or could reasonably be **linked**, directly or indirectly, with a particular consumer or household";[73] and includes within its definition of 'unique identifier': "a persistent identifier that can be used **to recognize** a consumer, a family, or a device that is linked to a consumer or family, over time and across different services, including, but not limited to, a device identifier" (emphasis added).[74]
- The 2019 Nigerian data protection regulation explicitly defines 'identifiable information' to include "information that can be used on its own or with other information to **identify, contact, or locate** a single person, or to identify an individual in a context" within its definition of 'personal data' (emphasis added).[75]
- The 2019 international standard in Privacy Information Management, ISO 27701, incorporates data which could be **"directly or indirectly linked"** to an individual, *regardless* of whether the individual

70 Dr Monique Mann and Professor Tobias Matzner, "Challenging algorithmic profiling: the limits of data protection and anti-discrimination in responding to emergent discrimination", Brussels Privacy Hub, *Working Paper*, Vol 6 (18), January 2020; available at https://brusselsprivacyhub.eu/publications/wp618.html

71 Dr Katharine Kemp, "Submission in Response to the Australian Competition and Consumer Commission Ad Tech Inquiry Issues Paper", 26 April 2020; available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3587239

72 Greenleaf & Livingston, "China's Personal Information Standard: The Long March to a Privacy Law", (2017) 150 *Privacy Laws & Business International Report*, pp. 25□28; available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3128593

73 CCPA section 1798.140(o)(1)

74 CCPA section 1798.140(o)(1)(x)

75 Nigeria Data Protection Regulation 2019; available at https://nitda.gov.ng/wp-content/uploads/2019/01/NigeriaDataProtectionRegulation.pdf

can be identified, within its definition of 'personally identifiable information' (emphasis added).[76]

4.2 Proposed new definition

So as to enable clarity and consistency in the application of data privacy law, and to protect against the potential privacy harms enabled by individuation, I propose that the data privacy laws which turn on identifiability as a threshold issue should be reformed to incorporate an expansive definition for the word 'identifiable', or explicitly add an alternative to identifiability.

Likewise, legislators considering the drafting of new data privacy laws should start with an expansive notion of 'personal data', so as to explicitly incorporate individuation as well as identifiability.

I propose ensuring that the definition of 'personal data' (or its equivalent) incorporates both identifiability and individuation.

For example, 'personal data' could be defined as meaning:

> *information or an opinion about or relating to an individual who is*
>
> *(i) identified or identifiable; or*
>
> *(ii) able to be discerned or recognised as an individual distinct from others, regardless of whether their identity can be ascertained or verified.*

I further suggest that the second element ("able to be discerned or recognised as an individual distinct from others") should then be defined as including:

> *if the individual, or a device linked to the individual, could (whether online or offline) be surveilled, tracked or monitored; or located, contacted or targeted; or profiled in order to be subjected to any action, decision or intervention including the provision or withholding of information, content, advertisements or offers; or linked to other data which is about or relates to the individual.*

This proposal aims to further explicate the notion of identifiability by individuation, by going to the heart of the types of privacy harms which can occur even when the precise 'identity' of an individual is unknown by the perpetrator of the harm. Mentioning decisions in relation to the provision or withholding of information is intended to capture the curation and delivery of personalised content such as ads, price offers, news feeds, recommendations for related content, etc. Further, the definition should cover decisions to ***exclude*** people from seeing certain content as much as it covers decisions to target or include people: e.g. a decision ***not*** to show a particular job ad to people outside a certain age bracket, or who identify as (or who have been inferred as belonging to) a particular ethnicity or religion.

This proposed additional layer to the test for identifiability also aims to ensure that the scope of data privacy regulation does not over-reach into technologies which do ***not*** pose risks of privacy harms, such as the use of sessional or load-balancing cookies which are necessary to make a website work, but which do not then continue to track the user.

76 ISO/IEC 27701:2019, available at https://www.iso.org/standard/71670.html

The result of such a reform would be that the act of placing a tracking cookie on a person's connected device, or using similar technology such as device fingerprinting, and then collecting data about that person's online behaviour in a way which distinguished them from other individuals, in order to profile and then target that person (for example, in order to serve up an advertisement, or to determine what offers or pricing to show that person) will constitute the handling of personal data, such that the privacy principles apply to that conduct, notwithstanding that the advertiser and online ad broker could each claim not to know (or even be able to find out) 'the identity' of the person.

By more explicitly embedding the concept of individuation within the core definitional element of 'identifiability', data privacy laws can evolve in the same direction as other recent privacy instruments such as the CCPA and ISO 27701.

I further suggest a definition to the effect that "device" should be read expansively, and can include a vehicle such as a car, a mobile device such as a mobile phone, a wearable such as a FitBit, an implantable such as a pacemaker, or a household device such as a smart TV. However, unlike the CCPA, the definition should ***not*** be limited with reference to the connectivity of devices. A vehicle could be tracked via aerial surveillance, or a photograph of its number plate, without needing the vehicle to 'connect' to the internet or to any other device.

4.3 Resolving pragmatic issues

A critical consideration for any proposed reform to the definition of personal data is how it would work in practice.

Data privacy laws typically contain actionable data subject rights relating to access and correction. If an individual has not been 'identified' by an organisation, but has only been 'singled out', how could the individual verify themselves such as to exercise their access or correction rights against that organisation?

This practical dilemma is not entirely novel. In Australia the definition of 'personal information' includes information "whether the information or opinion is recorded in a material form or not". If information is not recorded in a material form – for example, it has only been observed by an employee of a regulated entity – then how can the individual seek access or correction? The answer is they can't. The access and correction principles in Australia already only work in practice in relation to a sub-set of personal information, namely personal information that ***has*** been recorded in a material form. (By contrast, other privacy principles still make sense with respect to unrecorded information; gossiping about what a client did or looks like, without anything ever being recorded, could still constitute an unauthorised disclosure under the Australian Privacy Act.)

This paper suggests that an expanded scope for definition of personal data should not be rejected just because data subject rights cannot be realised for ***some*** types of personal data – a pre-existing problem in any case. The privacy of the individuals affected by individuation is still worthy of protection.

Whether in relation to access or correction, or any additional GDPR-type data subject rights such as a right to erasure, a right to algorithmic transparency, etc, the management of data subject rights needs careful consideration in terms of the degree of identity verification needed from an applicant. Regulated entities should avoid a situation in which access rights could be weaponised by a perpetrator of family violence to impersonate their ex-partner and seek access to geolocation data; or by any other motivated

intruder seeking to find out personal data about their target. This is an existing issue in relation to access requests.

On the other hand, the drafting of the definition of personal data should not be so prescriptive about being able to identify or single out ***individuals*** that it provides entities an excuse not to comply with their privacy obligations at all. For example, a data privacy legal framework should not be frustrated by an electricity provider claiming smart energy meter data is not personal data (and thus they don't need to protect it at all) just because they cannot single out an ***individual's*** use of electricity from the rest of their ***household***. This position – that linking data to a household is sufficient to attract both legal obligations and legal protections – is already accepted in jurisdictions including Europe[77] and New Zealand.[78]

I caution against following the approach taken in the drafting of the CCPA, which by explicitly including 'households' (or, in some contexts, 'families') as well as individuals within the definition of personal information, has the effect of extending all privacy rights and protections available under that Act to households (or families) as well as individuals. This is causing significant difficulties in practice, both for regulated businesses and for the Attorney General of California who must draft regulations under the Act.[79] Further, consumers have complained that in order to assert their privacy rights under the CCPA, they are actually being forced to hand over ***more*** personal information about themselves to companies than those companies already had.[80]

This paper offers the following model by way of solution:

- Access and correct rights (and any other data subject rights) should only apply to ***recorded*** personal data, where the applicant is able to demonstrate to an appropriate degree of certainty (commensurate with the level of risk posed by an unauthorised use or disclosure of the data sought) that they are either (i) the individual to whom the data relates or is about, or (ii) a nominated account holder on behalf of a household account, or (iii) authorised to represent all members of a household.
- Both privacy principles and data subject rights should apply notwithstanding that an entity cannot discern or recognise an individual as distinct from other members of the same household.

77 Lee Bygrave, *Data protection law: approaching its rationale, logic and limits*, Kluwer Law International, 2002, part 2.4.1.

78 See the 2010 inquiry into Google StreetView, available at https://www.privacy.org.nz/news-and-publications/commissioner-inquiries/google-s-collection-of-wifi-information-during-street-view-filming/ ; and Case Note 251185 [2015] NZ PrivCmr 3; available at https://www.privacy.org.nz/news-and-publications/case-notes-and-court-decisions/case-note-251185-2015-nz-privcmr-3-use-of-smart-meters-by-utility-companies/ ; and *R v Alsford* [2017] NZSC 42 at [30]; available at https://www.courtsofnz.govt.nz/assets/cases/2017/d2lj.pdf

79 For example, how can a 'household' exercise its rights of access/correction (or other rights under the CCPA such as deletion, or opting out, aka 'do not sell')? How is a household to verify its *bona fides* to a business which is expected to provide the household with access to (or deletion of) the 'personal information' held about the household? What if one family member uses their access to 'household' data to harm another family member? If notice is to be provided about a collection of personal information, how should the notice address a household? The draft regulations provide: "'Household' means a person or group of people who: (1) reside at the same address, (2) share a common device or the same service provided by a business, and (3) are identified by the business as sharing the same group account or unique identifier"; see Text of Modified Regulations [Clean Version] Title 11. Law Division 1. Attorney General Chapter 20. California Consumer Privacy Act Regulations Proposed Text Of Regulations, part 999.301(k); available at https://www.oag.ca.gov/sites/all/files/agweb/pdfs/privacy/ccpa-text-of-mod-clean-020720.pdf Yet confusingly, the CCPA separately refers to unique identifiers which are linked to, or can be used to recognise, "a consumer or family" (rather than household); family is then defined to mean minor children and their parent/guardian. Neither the Act nor the proposed Regulations make any attempt to reconcile these two competing groupings of individuals.

80 Alistair Barr, "Come on a trip into the new privacy circle of hell", *Bloomberg*, 9 January 2020; available at https://www.bloomberg.com/news/newsletters/2020-01-09/come-on-a-trip-into-the-new-privacy-circle-of-hell

- Provisions relating to making and responding to a complaint, the regulatory or supervisory authority's powers, data breach notification schemes, and all other provisions relating to enforcement and remedies, should apply notwithstanding that an individual complainant cannot demonstrate that the data at issue relates solely to themselves, as distinct from other members of the same household.

Conclusion

In this paper I have argued that in order to offer protection from privacy harms, data privacy laws need to recognise that in a digital environment, 'not identifiable' is no longer an effective proxy for 'will suffer no privacy harm'. Data privacy laws must anticipate the harms that can arise via individuation, or 'singling out' ***without identification***, as well.

I propose the word ***individuation*** to refer to the ability to disambiguate or 'single out' a person in the crowd, such that they could, at an individual level, be tracked, profiled, targeted, contacted, or subject to a decision or action which impacts upon them - even if that individual's 'identity' is not known (or knowable).

By more explicitly embedding the concept of individuation within the definition of 'personal data', data privacy laws around the world can be modernised to reflect the reality of our digital lives, and to protect against digital harms impacting on privacy and autonomy, such as social and market exclusion, discrimination, and manipulation of prices, emotions, and voting intentions, as well as tracking which can facilitate physical harms.

This paper offers the following as a six-step path forward for legislators, whether drafting a new statute or updating an existing data privacy law:

1. Ensure that the definition of 'personal data' (or its equivalent) incorporates information or opinion about or relating to an individual who is:

   (i) identified or identifiable; or

   (ii) able to be discerned or recognised as an individual distinct from others, regardless of whether their identity can be ascertained or verified.

2. Define the second element ("able to be discerned or recognised as an individual distinct from others") as including:

   if the individual, or a device linked to the individual, could (whether online or offline) be surveilled, tracked or monitored; or located, contacted or targeted; or profiled in order to be subjected to any action, decision or intervention including the provision or withholding of information, content, advertisements or offers; or linked to other data which is about or relates to the individual.

3. Define 'device' expansively, to include both connected and not-connected devices, including a vehicle such as a car, a mobile device such as a mobile phone, a wearable such as a FitBit, an implantable such as a pacemaker, or a household device such as a smart TV.

4. Access and correct rights (and any other data subject rights) should only apply to ***recorded*** personal data, where the applicant is able to demonstrate to an appropriate degree of certainty (commensurate with the level of risk posed by an unauthorised use or disclosure of the data sought) that they are either (i) the individual to whom the data relates or is about, or (ii) a nominated account holder on behalf of a household account, or (iii) authorised to represent all members of a household.

5. Both privacy principles and data subject rights should apply notwithstanding that an entity cannot discern or recognise an individual as distinct from other members of the same household.

and

6. Provisions relating to making and responding to a complaint, the regulatory or supervisory authority's powers, data breach notification schemes, and all other provisions relating to enforcement and remedies, should apply notwithstanding that an individual complainant cannot demonstrate that the data at issue relates solely to themselves, as distinct from other members of the same household.

The Brussels Privacy Hub Working Papers series

N°1 "The data protection regime applying to the inter-agency cooperation and future architecture of the EU criminal justice and law enforcement area" (November 2014) by Paul De Hert and Vagelis Papakonstantinou (35 pages)

N°2 "The new cloud computing ISO/IEC 27018 standard through the lens of the EU legislation on data protection" (November 2014) by Paul de Hert, Vagelis Papakonstantinou, Irene Kamara (25 pages)

N°3 "Towards efficient cooperation between supervisory authorities in the area of data privacy law" (October 2015) by Dariusz Kloza, Antonella Galetta (24 pages)

N°4 "The data protection regime in China" (November 2015) by Paul De Hert and Vagelis Papakonstantinou (30 pages)

N°5 "The right to privacy and personal data protection in Brazil: time for internet privacy rights?" (February 2016) by Vinícius Borges Fortes (23 pages)

N°6 "Permissions and Prohibitions in Data Protection Jurisdiction" (May 2016) by Mistale Taylor (25 pages)

N°7 "Structure and Enforcement of Data Privacy Law in South Korea" (October 2016) by Haksoo Ko, John Leitner, Eunsoo Kim and Jong-Gu Jung (20 pages)

N°8 "The "Right to be Forgotten" and Search Engine Liability" (December 2016) by Hiroshi Miyashita (15 pages)

N°9 "European Human Rights, Criminal Surveillance, and Intelligence Surveillance: Towards "Good Enough" Oversight, Preferably but Not Necessarily by Judges" (March 2017) by Gianclaudio Malgieri and Paul De Hert (25 pages)

N°10 "Does Technology Drive Law? The Dilemma of Technological Exceptionalism in Cyberlaw" (July 2017) by Meg Leta Jones, JD, PhD (31 pages)

N°11 "The Microsoft Ireland case and the cyberspace sovereignty trilemma. Post-territorial technologies and companies question territorial state sovereignty and regulatory state monopolies" (July 2018) by Paul De Hert and Johannes Thumfart (27 pages)

N°12 "Understanding the balancing act behind the legitimate interest of the controller ground: a pragmatic approach" (August 2018) by Irene Kamara and Paul De Hert (35 pages)

N°13 "Big data analytics by telecommunications operators and the draft ePrivacy Regulation" (September 2018) by Vagelis Papakonstantinou and Paul de Hert (13 pages)

N°14 "Enforcement in Indonesia Data Privacy Laws: The 2018 Facebook-Cambridge Analytica scandal as a case study" (October 2018) by Anbar Jayadi (21 pages)

N°15 "Belgium, Courts, Privacy and Data Protection. An inventory of Belgian case law from the pre-GDPR regime (1995-2015)." (January 2019) by Paul De Hert (34 pages)

N°16 Big data analytics in electronic communications: A reality in need of granular regulation (even if this includes an *interim* period of no regulation at all) (June 2019) by Vagelis Papakonstantinou and Paul de Hert (25 pages)

N°17 Data Localisation: Deconstructing myths and suggesting a workable model for the future. The cases of China and the EU (September 2019) by Author: Yanqing Hong, Senior Fellow, Law and Development Institute, Peking University of China, Edited by Vagelis Papakonstantinou, Brussels Privacy Hub (31 pages)

N°18 Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination (January 2020) by Dr Monique Mann and Professor Tobias Matzner (18 pages)

N°19 Access to the Internet in the EU: a policy priority, a fundamental, a human right or a concern for eGovernment? (February 2020) by Lina Jasmontaite and Paul de Hert (23 pages)

N°20 The Proposed ePrivacy Regulation: The Commission's and the Parliament's Drafts at a Crossroads? (March 2020) by Elena Gil González, Paul De Hert & Vagelis Papakonstantinou (31 pages)

N°21 Article 8 ECHR compliant and foreseeable surveillance: the ECtHR's expanded legality requirement copied by the CJEU. A discussion of European surveillance case law (April 2020) by Paul De Hert & Gianclaudio Malgieri (42 pages)

N°22 The "Ethification" of Privacy and Data Protection Law in the European Union. The Case of Artificial Intelligence (May 2020) by Niels van Dijk and Simone Casiraghi (23 pages)

N°23 Logic and Key Points of China's Cybersecurity Review Measures (June 2020) by Yanqing Hong, Senior Fellow, Law and Development Institute, Peking University of China, Edited by Vagelis Papakonstantinou, Brussels Privacy Hub (9 pages)

N°24 Individuation: re-imagining data privacy laws to protect against digital harms (July 2020) by Anna Johnston (22 pages)

The Brussels Privacy Hub Working Papers series

The Brussels Privacy Hub Working Papers are intended to circulate research in progress for comment and discussion. The Working Papers focus on all areas of data protection and privacy research and can contain empirical research on privacy issues and analytical work on privacy governance and regulation in the EU; global flows of data; reconciling law enforcement and privacy interests; privacy challenges posed by new technologies; comparative research on privacy in different regions; jurisprudential issues of privacy protection; and many others

Available at www.brusselsprivacyhub.org/publications.html

Contact: info@brusselsprivacyhub.eu

BRUSSELS PRIVACY HUB