The GDPR and the Internet of Things: A Three-Step Transparency Model

Sandra Wachter[1, 2]

[1]*Oxford Internet Institute, University of Oxford, 1 St Giles, Oxford, OX1 3JS, United Kingdom;*
[2]*The Alan Turing Institute, British Library, 96 Euston Rd, London, NW1 2DB, United Kingdom*

Corresponding author:
Dr. Sandra Wachter
Oxford Internet Institute
University of Oxford
1 St. Giles
Oxford, OX1 3JS
United Kingdom
sandra.wachter@oii.ox.ac.uk

Funding

This article is a deliverable of the Privacy-Enhancing and Identification-Enabling Solutions for IoT (PEIESI) project, part of the PETRAS Internet of Things research hub. PETRAS is funded by the Engineering and Physical Sciences Research Council (EPSRC), grant agreement no. EP/N023013/1.

Conflicts of interest

The author declares no actual or potential conflicts of interests. No financial interests or benefits have arisen from the direct application of this research.

Acknowledgements

The author is indebted to Dr. Mariarosaria Taddeo and Dr. Brent Mittelstadt of the University of Oxford, and the 'Ethics, Privacy, and Trust in IoT' workshop participants who provided invaluable feedback during preparation of the manuscript and improved the quality of the work greatly. The author would also like to thank the EPSRC for the funding provided to the PETRAS consortium which made preparation of this article possible. Finally, the author would like to thank the reviewers of the paper. The final version of this manuscript benefited greatly from their thoughtful comments and insightful feedback.

GDPR and the Internet of Things: Guidelines to Protect Users' Identity and Privacy

The Internet of Things (IoT) requires pervasive collection and linkage of user data to provide personalised experiences based on potentially invasive inferences. Consistent identification of users and devices is necessary for this functionality, which poses risks to user privacy. The General Data Protection Regulation (GDPR) contains numerous provisions relevant to these risks, which may nonetheless be insufficient to ensure a fair balance between users' and developers' interests. A three-step transparency model is described based on known privacy risks of the IoT, the GDPR's governing principles, and weaknesses in its relevant provisions. Eleven ethical guidelines are proposed for IoT developers and data controllers on how information about the functionality of the IoT should be shared with users above the GDPR's legally binding requirements. Two use cases demonstrate how the guidelines apply in practice: IoT in public spaces and connected cities, and connected cars.

Keywords: Data protection, Ethics, Privacy, Internet of things, Profiling

1 Introduction

The 'Internet of Things' (IoT) is a rapidly growing technology sector. In the EU, development and adoption of the IoT can be seen in areas such as health and wellness,[1] utilities,[2] urban planning and management,[3] logistics and supply chain management,[4] agriculture, and commerce.[5] Vast amounts of personal and usage data are now collected and shared by IoT devices and services.

A defining characteristic of the IoT is pervasive, often opaque collection and seamless linkage of user data to provide personalised experiences.[6] To enable this functionality, IoT

[1] Farzad Khodadadi, Amir Vahid Dastjerdi and Rajkumar Buyya, 'Internet of Things: An Overview' [2017] arXiv preprint arXiv:1703.06409 <https://arxiv.org/abs/1703.06409> accessed 30 June 2017; F Gonçalves and others, 'Security Architecture for Mobile E-Health Applications in Medication Control', *2013 21st International Conference on Software, Telecommunications and Computer Networks - (SoftCOM 2013)* (2013); Cisco, 'Securing the Internet of Things: A Proposed Framework' (2016) <http://www.cisco.com/c/en/us/about/security-center/secure-iot-proposed-framework.html> accessed 6 July 2017.

[2] S Sicari and others, 'Security, Privacy and Trust in Internet of Things: The Road Ahead' (2015) 76 Computer Networks 146; Khodadadi, Dastjerdi and Buyya (n 1).

[3] Sicari and others (n 2); Khodadadi, Dastjerdi and Buyya (n 1).

[4] Sicari and others (n 2); C Yuqiang, G Jianlan and H Xuanzi, 'The Research of Internet of Things' Supporting Technologies Which Face the Logistics Industry', *2010 International Conference on Computational Intelligence and Security* (2010); L Weiss Ferreira Chaves and C Decker, 'A Survey on Organic Smart Labels for the Internet-of-Things', *2010 Seventh International Conference on Networked Sensing Systems (INSS)* (2010).

[5] Sicari and others (n 2).

[6] Sandra Wachter, 'Normative Challenges of Identification in the Internet of Things: Privacy, Profiling, Discrimination, and the GDPR' (Social Science Research Network 2017) SSRN Scholarly Paper ID 3083554 <https://papers.ssrn.com/abstract=3083554> accessed 12 December 2017; Omer Tene and Jules Polonetsky, 'Big Data for All: Privacy and User Control in the Age of Analytics' [2013] Nw.J. Tech. & Intell. Prop.

devices and services must be connected and share data about users' interactions with multiple nodes in the network. Consistent identification of users and devices across the network is likewise necessary.

These features of the IoT, which create numerous privacy risks, are frequently designed to go unnoticed by users in order to provide a 'seamless' experience.[7] Potentially invasive inferences can be drawn from linked datasets, including data generated through usage of connected devices and services.[8] Inferential analytics can drive personalised, potentially discriminatory decision-making.[9] The impossibility of anonymising data,[10] weak cybersecurity standards,[11] and the opaque operation of many IoT devices and services further exacerbate these privacy risks, and users' awareness of them.[12] A fundamental tension exists between the seamless and non-transparent nature of the IoT, and the need to keep users informed and in control of collection and processing of their personal data to protect against privacy threats.

In Europe, risks of profiling and invasive inferential analytics enabled by pervasive data collection and seamless linkage are reflected in the regulatory landscape.[13] The General Data Protection Regulation (GDPR) contains numerous provisions relevant to the risks posed by identification technologies. However, the strict legal requirements defined in the Articles of the GDPR may be insufficient to ensure a fair balance is struck between user's interests in privacy and the interests of IoT developers and data controllers.

To address this gap, this paper proposes a three-step transparency model based on known privacy risks of the IoT, weaknesses in relevant legally binding provisions in the GDPR, and the GDPR's governing principles. Eleven guidelines aimed at IoT developers and data

<http://heinonlinebackup.com/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/nwteintp11&section=20> accessed 2 October 2014.

[7] Wachter (n 6); Scott R Peppet, 'Regulating the Internet of Things: First Steps toward Managing Discrimination, Privacy, Security and Consent' (2014) 93 Tex. L. Rev. 85.

[8] Sarah Johanna Eskens, 'Profiling the European Citizen in the Internet of Things: How Will the General Data Protection Regulation Apply to This Form of Personal Data Processing, and How Should It?' (Social Science Research Network 2016) SSRN Scholarly Paper ID 2752010 <https://papers.ssrn.com/abstract=2752010> accessed 8 July 2017; W Kuan Hon, Christopher Millard and Jatinder Singh, 'Twenty Legal Considerations for Clouds of Things' <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2716966> accessed 8 July 2017.

[9] Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' [2018] Columbia Business Law Review, forthcoming (2019) 84; Solon Barocas and Andrew D Selbst, 'Big Data's Disparate Impact' (2016) 104 California Law Review.

[10] Paul Ohm, 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1450006> accessed 28 June 2017.

[11] R Roman, P Najera and J Lopez, 'Securing the Internet of Things' (2011) 44 Computer 51.

[12] Wachter (n 6).

[13] Sandra Wachter, 'Privacy: Primus Inter Pares — Privacy as a Precondition for Self-Development, Personal Fulfilment and the Free Enjoyment of Fundamental Human Rights' <https://papers.ssrn.com/abstract=2903514> accessed 14 September 2017.

controllers are described. These guidelines describe how information about the functionality of IoT devices and services should be shared with users to better inform their choices around data collection, processing, storage, and transfer. Specifically, user trust and acceptance can be enhanced by (a) openly describing the possible risks (e.g. discrimination) of IoT systems; (b) demonstrating the existence of robust obscurity tools that allow users to restrict inaccurate or unwanted inferential analytics and profiling; and (c) showing contingency plans are in place to mitigate privacy risks if systems are compromised. The guidelines describe ethically desirable requirements to be adhered to in addition to the GDPR's legally binding requirements. To demonstrate how the guidelines could apply in practice and alter the design choices and practices of IoT developers and data controllers, two use cases are considered: IoT in public spaces and connected cities, and connected cars. These cases reveal how application of the guidelines differs according to the types of space monitored (e.g. public, private, and mixed) and people monitored (e.g. unidentified, known, and incidental 'users').

2 Background

A recent literature review – highly inspired by Peppet's harm taxonomy of the IoT[14] – highlighted a tension between privacy and identification in the IoT.[15] The review identified four primary challenges relevant to the design and regulation of identification technologies in the IoT: (1) linkage of user identities and records generated from IoT devices and services, which can lead to potentially invasive profiling, inferences, and discrimination; (2) disclosure of sensitive information to other IoT users and data controllers that the data subject would otherwise prefer to keep confidential, and inhibiting user's control over such disclosures; (3) creation of information or inferences about the user, which could not have been predicted when the user set access policies or chose to use the device/service; and (4) limitations on user oversight and transparency in management of identity and profiling, which can facilitate breaches of privacy and undermine trust.[16] Addressing these challenges proactively requires legal and ethical alignment of IoT design choices, business practices, and regulatory provisions.

Given the necessity of pervasive collection and seamless linkage of personal data to enable identification in the IoT, data protection and privacy laws are particularly relevant.[17] Data protection law explicitly deals with the question of how to balance privacy with the free

[14] Peppet (n 7).
[15] Wachter (n 6).
[16] ibid.
[17] Hon, Millard and Singh (n 8).

flow of data and other business interests. In Europe, the legal landscape recently experienced a significant change with the GDPR that came into force on 25th May 2018.[18] The GDPR aims to create a harmonised data protection standard across the EU in order to strike a balance between the free flow of data and the fundamental interests of data subjects (e.g. privacy). As the IoT collects, processes, and shares substantial volumes and varieties of personal data, the GDPR must be treated as a key governance framework for the design and deployment of IoT systems.

The GDPR has introduced new governing data protection principles (Article 5 and 25) and standards with which developers and data controllers for IoT devices and services must comply. In the context of privacy risks of identification technologies in the IoT, the GDPR is a particularly relevant legal framework due to its applicability across all sectors that process personal data, and its wide geographical impact.[19] Standards relating to informed consent, notification duties, privacy by design and privacy by default, data protection impact assessment, algorithmic transparency, automated decision-making, and profiling now apply across Europe and beyond, and may help address the tension between privacy and identification in the IoT.

The aforementioned review critically examined specific provisions of the GDPR relevant to privacy and identification in the IoT, including transparency (Article 5), data storage, access, rectification, and deletion (Articles 5, 15-17), informed consent (Article 7), notification duties (Articles 13-14 and 33-34), automated decision-making and profiling (Articles 21-22), privacy by design and privacy by default (Article 25), cybersecurity (Articles 33-34), and data protection impact assessment (Article 35-36).

The review concluded that several of these provisions urgently require further specification and implementation into the design and deployment of IoT technologies to minimise the impact of IoT profiling and identification technologies on user privacy. Key concepts are left vague or undefined in the GDPR. This creates ambiguity in how to balance

[18] Note that the ePrivacy Regulation – a lex specialis to the GDPR - is set to come into force in 2018 as well, however the scope of this framework goes beyond the scope of this paper.

[19] As Wachter (n 6). argues, the jurisdiction of the GDPR is wide in both technological and international terms: "The GDPR applies to all data controllers using personal data. The framework will apply to all IoT devices (e.g. smart or automobile sensors) and all sectors (e.g. health, transport) that generate or process personal data. Further, the GDPR will apply not only to Member States, but also by extension to certain data controllers in third countries. In Article 3 the GDPR defines the scope of the framework and states that data controllers (e.g. US companies) without establishment in the EU are nonetheless subject to the GDPR if they offer products and services involving personal data processing in the European Union."

the interests of data subjects in privacy, and the interests of data controllers in identification and providing linked-up IoT services. For example, requirements to notify data subjects in case of data breaches (Article 34) apply only to breaches likely to impose a 'high risk to the rights and freedoms of natural persons.' Unfortunately, 'high risk' is left undefined, meaning it is unclear which sectors or specific data types are perceived as most concerning.

Elsewhere, limitations are imposed on the scope of protections that must be supplied by data controllers, which minimises the protection they offer against privacy invasive identification, inferential analytics, and profiling. For example, Article 22, which addresses automated decision-making and profiling, limits the definition of 'automated individual decision-making' to decisions affecting data subjects 'based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.' As several commentators have noted,[20] this definition includes undefined terminology (i.e. 'solely automated', 'legal or similarly significant effects') that may introduce a loophole in which nominal human involvement in a computer-driven decision-making process renders the provisions inapplicable.

Data controllers in the IoT thus face a double challenge: while operating systems designed to work seamlessly and in the background, they must nonetheless keep users informed and in control of their data according to poorly defined data protection standards. Highlighting this challenge, several specific points of conflict and difficult vagueness between GDPR provisions and identification in the IoT were also noted in the aforementioned review.[21] First, IoT devices and services are often characterised by 'data maximalism', or the excessive collection, storage, and sharing of personal data on the basis that it may prove useful in the future. This tendency directly conflicts with calls for data minimalism or purpose limitation

[20] Sandra Wachter, Brent Mittelstadt and Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' [2017] International Data Privacy Law <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2903469>; Sandra Wachter, Brent Mittelstadt and Chris Russell, 'Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR' (2018) forthcoming Harvard Journal of Law & Technology; Isak Mendoza and Lee A Bygrave, 'The Right Not to Be Subject to Automated Decisions Based on Profiling' in Tatiani Synodinou and others (eds), *EU Internet Law: Regulation and Enforcement* (Springer 2017) <https://papers.ssrn.com/abstract=2964855> accessed 10 May 2017; Michael Veale and Lilian Edwards, 'Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling' (Social Science Research Network 2017) SSRN Scholarly Paper ID 3071679 <https://papers.ssrn.com/abstract=3071679> accessed 27 December 2017; Lilian Edwards and Michael Veale, 'Slave to the Algorithm? Why a "Right to an Explanation" Is Probably Not the Remedy You Are Looking For' (Social Science Research Network 2017) SSRN Scholarly Paper ID 2972855 <https://papers.ssrn.com/abstract=2972855> accessed 12 August 2017.

[21] For similar concerns see Tal Zarsky, 'Incompatible: The GDPR in the Age of Big Data' (2017) 47 Seton Hall Law Review <https://papers.ssrn.com/abstract=3022646> accessed 26 February 2018.

(Article 5(1)(b)), informed consent for specific and well-defined purposes (Article 7), and privacy by design (Article 25).

Second, complex inferential analytics used to profile users and provide personalised services can reveal unforeseen correlations and information about data subjects. This aspect of the IoT again conflicts with expectations that informed consent will be granted for specific and well-defined purposes (Article 7). Further, data controllers are expected under specific circumstances to conduct a data protection impact assessment (DPIA; Article 35) in which the potential risks of processing should be identified. The uncertain value of personal data generated and processed by IoT devices and services necessarily limits the scope of risks that can be foreseen, and thus reduces the protection actually offered by DPIAs.

Third, recognising this uncertainty, notification requirements imposed on data controllers (Articles 13-14) may be insufficient to deliver meaningful transparency that conveys to data subjects the complexity and uncertainty of using the IoT, and its associated data linkage, profiling, and inferential analytics. Data controllers may, for example, be allowed to communicate risks via generic templates or icons aimed at lay audiences, which poorly inform users about their subjective risks or loss of control over their identity.[22] Users' ability to make informed choices about which IoT applications to use, and how to manage the collection, processing, and transfer of the personal data essential to their functionality would be inhibited by such forms of disclosure.

Finally, it remains unclear how much protection data subjects' interests will receive when in conflict with the 'legitimate interests' of data controllers. In connection with the principle of transparency (Article 5), Articles 15 to 17 specify several rights for data subjects to exercise control over disclosures of personal data, and thus prevent invasions of privacy or discriminatory treatment fuelled by the IoT. These rights can, however, be overridden by the 'legitimate interests' of data controllers in some cases. Guidance to strike a fair balance between the interests of both parties is not offered by the GDPR.[23]

[22] Wachter, Mittelstadt and Russell (n 20).
[23] Wachter and Mittelstadt (n 9).Some guidance can be found here: Article 29 Data Protection Working Party, 'Opinion 06/2014 on the Notion of Legitimate Interests of the Data Controller under Article 7 of Directive 95/46/EC' 844/14/EN WP 217 <http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf> accessed 8 January 2018.

3 Principles and guidelines for transparency and trust in the IoT

Given the privacy risks of the IoT and the lack of clarity in key provisions of the GDPR relevant to the IoT, going forward data subjects may be exposed to devices and services that strike a legally compliant but ethically undesirable balance between privacy and identification. The GDPR may, however, provide alternative grounds to resolve tensions between privacy and identification in the IoT. In particular, the GDPR's governing principles of lawful data processing (Article 5) may provide grounding to resolve tensions between privacy and identifiability in the IoT. In Section 3.1, points of conflict between these principles and identifiability in the IoT are reviewed. An argument is then made in Section 3.2 for implementing extra-legal guidelines inspired by the GDPR's governing principles to resolve these tensions. Such guidelines can ensure an ethically desirable level of protection for the privacy and identity of users, extending beyond the GDPR's strict legal requirements.

3.1 Governing Principles of the GDPR

Several points of conflict between the GDPR's governing principles and identification in the IoT can be observed. The governing principles of the GDPR as defined in Article 5 are:

1. Lawfulness, fairness, and transparency (Article 5(1)a)

This trinity of principles describes data controllers' obligations to have legitimate grounds for processing of personal data. To ensure the lawfulness of the processing, transparency plays a key role. Data subjects should be aware of the processing purposes and should be provided with suitable notification and information regarding its scope. Even though fairness is not defined, the Article 29 Working Party and scholars believe that fairness links to awareness, meaning data subjects should be made aware of data processing.[24] This is especially relevant for IoT developers since devices often collect vast amounts of personal data, some of which can be considered sensitive (e.g. FitBit, health data).[25] The seamless implementation of these techniques can cause users to forget that their data is constantly being collected.[26]

2. Purpose limitation (Article 5(1)b)

[24] Eskens (n 8); Article 29 Data Protection Working Party, 'Opinion 8/2014 on the on Recent Developments on the Internet of Things' (2014) 14/EN WP 223 <http://www.dataprotection.ro/servlet/ViewDocument?id=1088> accessed 8 July 2017; Lee A Bygrave, *Data Protection Law: Approaching Its Rationale, Logic and Limits* (Kluwer Law Intl 2002) 59; Luiz Costa and Yves Poullet, 'Privacy and the Regulation of 2012' (2012) 28 Computer Law & Security Review 254, 256.

[25] Rosemary Jay, *Guide to the General Data Protection Regulation: A Companion to Data Protection Law and Practice* (4th Revised edition edition, Sweet & Maxwell 2017).

[26] Eskens (n 8) 37; Mark Weiser, 'The Computer for the 21st Century' (1991) 265 Scientific american 94; Neil Gershenfeld, *When Things Start to Think* (Macmillan 1999).

The principle of purpose limitation refers to the obligation of data controllers to only use the collected data for specific and well-defined purposes. The usage of collected data for other purposes has to be compatible with the initial one. Consent of the data subject or Member State laws can offer grounds to legitimise additional processing not related to the initial purpose.[27] This principle can pose difficulties for the IoT.[28] Very often vast amounts of data are collected for vague or broadly defined purposes.[29] Sensor fusion,[30] or the linkage[31] of existing but previously unconnected datasets, can offer new opportunities for data analytics that were not envisioned when the data were collected. Invasive and unpredictable inferential profiling is enabled by identification services that link devices and the data they collect.[32] Absent meaningful transparency addressing how user data is being used, these characteristics of the IoT substantially undermine users' capacities to protect their privacy and control their identity.

3. Data minimisation (Article 5(1)c)

Data controllers are required to only use data that are 'adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed.'[33] Data controllers must ensure that the collected data are necessary for their intended processing scope, and that excessive data are not collected beyond this scope. For the IoT, data controllers must establish that the data being collected is necessary to deliver their product or services. This principle challenges the typical 'data maximalism' of the IoT and Big Data analytics by extension, which require vast data collection and linkage for the personalisation of services (but not for the immediate functionality of a solitary device or service).[34]

4. Accuracy (Article 5(1)d)

[27] Jay (n 25) 83–89.
[28] Lokke Moerel and Corien Prins, 'Privacy for the Homo Digitalis: Proposal for a New Regulatory Framework for Data Protection in the Light of Big Data and the Internet of Things' [2016] Browser Download This Paper <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2784123> accessed 8 July 2017.
[29] Paul de Hert and Vagelis Papakonstantinou, 'The New General Data Protection Regulation: Still a Sound System for the Protection of Individuals?' (2016) 32 Computer Law & Security Review 179.
[30] Ohm (n 10) 118.
[31] Salvatore Ruggieri, Dino Pedreschi and Franco Turini, 'Data Mining for Discrimination Discovery' (2010) 4 ACM Transactions on Knowledge Discovery from Data (TKDD) 9; Tal Zarsky, 'The Trouble with Algorithmic Decisions An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making' (2016) 41 Science, Technology & Human Values 118.
[32] Wachter (n 6).
[33] REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) 2016 Art 5(1)(c), Recital 39.
[34] Wachter (n 6).

Data controllers are required to only store and use data that are accurate. Accuracy refers to the need for data to be correct and complete with regard 'to the purposes for which they are processed.' Incorrect data must be rectified or deleted without undue delay.[35] As a result, IoT developers face a significant challenge to curate and update their datasets to meet this requirement. Verification of user identity is critical to ensure accuracy, particularly when multiple people can potentially use the same device. Without verification, usage data from multiple users could be mistakenly recorded under a single user's profile, leading to inaccurate processing.

5. Storage limitation (Article 5(1)e)

The principle of storage limitation obligates data controllers to not store personal data for 'longer than is necessary for the purposes for which the personal data are processed.' In the IoT, the utility of stored data for the intended purpose of a particular product or service will need to be periodically re-assessed. Storage is also allowed without a link to a specific processing purpose when data 'will be processed solely for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes.' This principle can conflict with competing interests and rights of data subjects (e.g. right of access, right to be forgotten) or other obligations according to Member State laws that require longer or shorter periods of data storage (e.g. Art 23 GDPR addresses access to historical data for criminal investigations).[36]

6. Integrity and confidentiality (Article 5(1)f)

Data controllers are required to implement appropriate security mechanisms to guard against unlawful access, data breaches, data losses or leaks. For IoT developers, appropriate (cyber)security standards and mechanisms must be embedded in the design of devices and services. This requirement can be particularly challenging for technologies with simplistic functionality or low computational power (e.g. RFID or WiFi), which cannot support intensive mechanisms such as encryption.[37] The effectiveness of security mechanisms can quickly fade due to newly identified weaknesses or types of attacks. Integrity and confidentiality therefore

[35] REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (n 33).
[36] Hon, Millard and Singh (n 8); Roba Abbas, Katina Michael and MG Michael, 'Using a Social-Ethical Framework to Evaluate Location-Based Services in an Internet of Things World' (2015) 22 International Review of Information Ethics 42.
[37] Ohm (n 10).

appear to require long-term commitment by IoT developers to identify new threats, and patch their devices and services accordingly.[38]

7. Accountability (Article 5(2))

The principle of accountability should be achieved through three main duties derived from the six preceding principles.[39] First, data controllers are obligated to keep records of their data processing activities. Second, they should implement 'privacy by design' and 'privacy by default' mechanisms. Third, data controllers are required to undertake a data protection impact assessment for high-risk data processing. These provisions aim to ensure that data controllers take their obligations to respect all the fundamental principles seriously, and can demonstrate compliance as required.[40] Requirements for impact assessments specific to the IoT have not yet been derived from the aforementioned principles, but will need to grapple with a multitude of application- and sector-specific risks.

3.2 Extra-legal Guidelines for IoT Developers Using Identification Technologies

The seven governing principles of the GDPR are critical for balancing privacy, trust, and identifiability in the IoT. Profiling and subsequent unlawful discrimination cannot always be prevented. Protection of privacy and the resilience of systems against cyber-attacks similarly cannot always be guaranteed. Rather than only focusing on the untenable promise to guarantee privacy at all times, fostering user trust through transparency and honest communication of risks may be a better option. Openness and honesty about possible risks might be preferable to leading users to believe that their interests will be protected in all cases. Users require high quality, understandable, and sufficiently broad information to make an informed decision about whether to trust and ultimately adopt a system. Dialogue between developers and users is critical because IoT is seamless, often hidden, and can lead to unpredictable and opaque discrimination. Potential users are less likely to adopt the IoT if providers and applications are not perceived as trustworthy, meaning the anticipated benefits and efficiency promised by the IoT may not materialise if these risks are not taken seriously at an early stage.[41]

[38] Roman, Najera and Lopez (n 11).
[39] Jay (n 25) 169 ff.
[40] ibid.
[41] Wachter (n 6); Wachter, Mittelstadt and Russell (n 20); Wachter (n 13); Sicari and others (n 2); PN Mahalle and others, 'A Fuzzy Approach to Trust Based Access Control in Internet of Things', *Wireless VITAE 2013* (2013); M Nitti and others, 'A Subjective Model for Trustworthiness Evaluation in the Social Internet of Things', *2012 IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications - (PIMRC)* (2012).

The governing principles of the GDPR provide a robust foundation to propose supplementary guidelines that can help close the gap between the GDPR's explicit legal requirements and ethically desirable design and communication in the IoT. A higher level of protection is desirable for both IoT developers and data controllers on the one hand, and users on the other. Disclosing greater and more meaningful information than is strictly legally required regarding how the IoT collects and handles users' personal data, or in other words greater transparency on the supply side, encourages greater trust between users and suppliers.

Such extra-legal guidelines should not, however, only seek to achieve greater transparency in the IoT by addressing perceived gaps in data protection law. Enforcement of privacy as a prohibition on personal data processing is also essential. To understand the distinction between these aims, Paul de Hert and Serge Gutwirth's formulation of the right to data protection and right to privacy is instructive. According to this approach, data protection law is a 'tool of transparency', used to protect individuals from abuses of power and harmful data processing by more powerful actors. Article 8 of the Charter of the Fundamental Rights of the European Union enshrines the 'right to data protection', which grants individuals a right to have personal data processed fairly and on a legitimate legal basis. Data protection law primarily achieves this by describing protective conditions to be met when personal data is processed, enforced through informational obligations and individually enforceable rights. Prohibitions on processing are less common in data protection law; rather, the aim is to describe conditions for legitimate personal data processing, and to ensure sufficient information is available to data subjects to verify these conditions have been met.[42]

In contrast, the 'right to privacy' can be conceived of as a "tool of opacity," which provides data subjects with "stopping power" and sets "normative limits" on the power institutions that process personal data.[43] Article 7 of the Charter grants individuals rights to withhold and withdraw information about their private life, and more broadly prohibits undue interference in the privacy life of individuals by public authorities. Although not absolute, the right to privacy nevertheless actively prohibits invasive forms of interference in private life, including interference via personal data processing. Data protection feeds into the protection of privacy with conditions for processing, and ensures that sufficient information is made available to allow data subjects to effectively exercise informational self-determination.[44]

[42] Paul De Hert and Serge Gutwirth, 'Privacy, Data Protection and Law Enforcement. Opacity of the Individual and Transparency of Power' [2006] Privacy and the criminal law 61.

[43] ibid.

[44] ibid.

Transparency is thus a complementary and necessary set of protections, but an insufficient check on the power of data controllers by itself. It follows that suppliers and data controllers must take both data protection (or transparency) and privacy (or opacity) seriously in the design and governance of the IoT.

In the following section, eleven extra-legal guidelines are proposed to restore a fair balance between transparency (or the right to data protection), opacity (or the right to privacy), and identifiability in the IoT. Together, these guidelines form a three-step transparency model which can help data subjects comprehend the actual privacy risks of profiling and identification technologies in the IoT, and thus be better placed to control disclosures of personal data and modifications to their identity.[45]

4 The three-step transparency model

As outlined in Section 2, the legally binding provisions of the GDPR provide insufficient protection to the privacy and identity of users. Extra-legal guidelines following the spirit of the law (e.g. Article 5's governing principles) are therefore desirable to help users better understand the extent and risks of collection, processing, and transfer of personal data by IoT devices and services. With this information, users will be better placed to make more informed choices regarding their adoption of IoT devices and services, and to more effectively manage how their personal data is collected, processed, and transferred.

To these ends, in this section a three-step transparency model is proposed which describes ethical ideals for transparency and disclosures by data controllers and IoT suppliers, and describes several tools of opacity to help users hide and control personal information. The model is intended to inform IoT developers and data controllers about how to mitigate some of the risks related to the IoT, and to comply with the spirit of the GDPR's guiding principles when its legally binding provisions offer insufficient protection to data subjects. The proposals made here meet calls in EU policy to define principles and guidelines for IoT devices.[46]

[45] Wachter (n 13); Mireille Hildebrandt, 'Profiling and the Identity of the European Citizen' [2008] Profiling the European citizen 303; Mireille Hildebrandt, *Smart Technologies and the End (s) of Law: Novel Entanglements of Law and Technology* (Edward Elgar Publishing 2015); Parikshit Mahalle and others, 'Identity Management Framework towards Internet of Things (IoT): Roadmap and Key Challenges' [2010] Recent Trends in Network Security and Applications 430; Alessandro Mantelero, 'The Future of Consumer Data Protection in the EU Re-Thinking the "Notice and Consent" Paradigm in the New Era of Predictive Analytics' (2014) 30 Computer Law & Security Review 643.
[46] Wachter (n 6).

The three-step transparency model consists of eleven extra-legal guidelines responsive to three areas: (1) the GDPR's governing principles (Article 5); (2) ambiguities and ethically undesirable limitations in provisions of the GDPR relevant to the IoT; and (3) known risks to privacy owing to profiling and identification in the IoT. The model argues that it is both untenable and misleading for IoT suppliers and data controllers to make absolute promises regarding the protection of user privacy. To achieve meaningful data protection via transparency and user privacy through opacity, and thus to enhance user trust, data controllers should (1) describe the possible risks (e.g. discrimination) of IoT systems openly (e.g. notification, data protection impact assessment, privacy policies); (2) show what kind of mechanisms are in place to limit inaccurate or unwanted predictions and assumptions, and consequently discrimination based on profiling (e.g. agile consent models, accurate prediction models, right of access, ethical sharing practises with third parties, opt-out options from profiling, algorithmic transparency and anti-discrimination tools in automated decision-making, and profiling); and (3) show transparent contingency plans to mitigate risks (discrimination) if the system is compromised (e.g. cyber risks, notification of data breaches, privacy enhancing technologies).

4.1 First step: Transparent information about possible risks

1. Data Protection Impact Assessment (DPIA)

Whenever 'systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling' and new technologies for data processing are used, a DPIA will be mandatory if the processing is 'likely to result in a high risk to the rights and freedoms of natural person'. Due to the growing importance of the IoT and the associated risks for privacy, a DPIA will be mandatory for most IoT devices. IoT developers will need to assess the possible risks of their devices. If their assessment indicates a high privacy risk, prior consultation of a supervisory authority will be mandatory.

Even though the Article 29 Working Party has issued guidelines stating that the DPIA should be (at least in parts) publicly available and should be 'continuously reviewed and regularly re-assessed'[47], their recommendation is not legally binding. The GDPR does not address this issue. However, it is recommended to consider to publish the results and the

[47] Article 29 Data Protection Working Party, 'Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing Is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679' (2017) 17/EN WP 248.

methods of the DPIA and to regularly review the document to help affected data, regulators, and national supervisory bodies effectively address and respond to risks as they materialise.

This type of iterative transparency will help data subjects to better understand the possible risks of their usage of an IoT product or service, and thus to make more informed choices when consenting to data processing, enhancing their ability to protect their privacy through opacity.[48] Greater communication of risks can help increase trust in IoT devices.[49] Further, even if a DPIA is not legally mandated, IoT developers should consider evaluating their technologies nonetheless. In cases where a DPIA is deemed unnecessary, a public statement of the reasons behind this decision can have a similar effect. This would help increase trust in IoT devices, as users can see that data controllers take their privacy seriously, evaluated possible risks carefully, and went beyond what is legally mandated to ensure privacy.

2. Communicate precisely, but elaborate when needed

Article 12 aims to ensure transparent information and communication to enable data subjects to exercise their rights as defined in the GDPR. The language used should be in a 'concise, transparent, intelligible and easily accessible form, using clear and plain language',[50] indicating that the imagined audience is a lay person.[51] These requirements are even more important when children are addressed (Art 12 (1)).

While it is intuitively preferable to communicate with data subjects with concise and easily understandable language for the sake of simplicity and to avoid confusion, this approach also limits the quality of the information being conveyed. Possible negative consequences of data collection and processing, including leaks owing to hacking, invasive inferences due to sensor fusion, and the broader predictive and inferential power of big data analytics can be difficult to communicate with easily understandable language. More elaborate communication may be necessary when disclosing uncertain but high impact risks, for example the knock-on effects of a data breach or identification of users by a third party. Greater detail regarding the nature and likelihood of high impact risks is essential to ensure IoT users have sufficient

[48] De Hert and Gutwirth (n 42).
[49] Wachter (n 6); Wachter (n 13); Wachter, Mittelstadt and Russell (n 20); Mahalle and others (n 41); Nitti and others (n 41); Sicari and others (n 2).
[50] REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (n 33).
[51] Frank, 'Art 12 Transparente Information, Kommuikation' in Peter Gola and Carolyn Eichler (eds), *Datenschutz-Grundverordnung VO (EU) 2016/679: Kommentar* (CH Beck 2017) Art 12 Rn 17-23.

information to make informed choices regarding continued usage and data management, or to take additional steps to protect their privacy following a breach.

3. Icons might not always be the best tool for communication

To achieve lawfulness, fairness, and transparency, data subject awareness is key. The GDPR implements new obligations related to transparency. This is reflected in Articles 13-14, which create notification duties for data controllers. Amongst other things, data controllers have to inform data subjects about intended data processing purposes, contact details of the data controller, the recipients of the subject's personal data, the period for which the personal data will be stored, the usage of profiling and the right to object to it (Articles 13(2)(b) and 14(2)(c)), and the existence of automated decision-making, including profiling (Articles 13(2)(f) and 14 (2) g)). It is reflected in Art 12(7) that the information about intended processing purposes referred to in Articles 13-14 can be conveyed using standardised icons alongside short texts.[52] Most IoT devices have small screens that make reading policy statements harder, which can be problematic if freely given and informed consent is legally required for processing.[53] This would be the case when an IoT device or service is collecting or handling sensitive data, or when an alternative legal basis for processing (e.g. Member State law, the 'legitimate interests' of the data controller) is unavailable.

Similar to the concerns above, since the information provided aims to inform data subjects about what will happen to their data, and to enable them to make an informed decision about engaging in those processes, standardised icons and short descriptions may prove insufficient. In particular, the requirement to inform data subjects about the logic involved in automated decision-making (including profiling) will be challenging in this regard due to the inherent opacity and complexity of algorithmic systems.[54]

Even though the simplicity and standardised communication offered by icons are desirable, their power to inform is limited, even if accompanied with short descriptive text. Additional information about the functionality of systems being used, particularly in the case of complex algorithms and machine learning tools, should be provided for users who want to learn more, especially since the opacity and inscrutability of AI based systems offers a great source for discrimination. Meeting this higher ethical standard for transparency will help ensure

[52] Article 29 Data Protection Working Party, 'Guidelines on Transparency under Regulation 2016/679' (2017) 17/EN WP 260.
[53] Peppet (n 7) 140.
[54] Brent Mittelstadt and others, 'The Ethics of Algorithms: Mapping the Debate' [2016] Big Data & Society.

that users are not unduly giving away personal data and insights they might otherwise prefer remain hidden.

4. Privacy should not be the foe of transparency

In order to guarantee trust and awareness of data processing, the GDPR not only requires data controllers to notify data subjects about intended processing purposes (Articles 13-14), but also allows data subjects to request more or less the same information at any time under the right of access (Article 15). The right of access empowers data subjects to independently manage privacy without relying on data controllers to provide appropriate and timely information.[55] Information regarding the scope and purpose of data processing can be obtained via the right of access, without which other rights such as rectification (Art 16), erasure (Art 17), or objection to processing (Art 21) are impossible to exercise effectively.

At the same time Art 15(4) and Recital 63 allow data controllers to limit the requested information based on conflicts with the rights and freedoms of others. These freedoms include the privacy rights of other data subjects or interests of data controllers such as trade secrets and intellectual property rights.[56] The GDPR calls for a fair balance between the interests of individuals, other data subjects, and data controllers. A balance is thus required between supplier side transparency that enables user privacy (or opacity) by definition, and the privacy (or opacity) interests of other data subjects. Simply put, users do not have an absolute right to supplier transparency if such disclosures risk exposing the private information of other data subjects.

Finding this balance will prove very challenging in cases where information about profiling and automated decision-making is requested. The profiles used are usually built on data from reference groups (e.g. personal data of other users). Group privacy rights are not sufficiently acknowledged in current data protection law, which focuses on the individual data subject rather than the collective.[57] This fact could be used as a loophole to not disclose information about profiling, since it could be claimed that this information infringes other data subjects' privacy. Concerns with the privacy of others should not be misused to prevent access

[55] Eugen Ehmann, 'Art 15 Rechte der betroffenen Person' in Eugen Ehmann, Martin Selmayr and Jan Philipp Albrecht (eds), *DS-GVO: Datenschutz-Grundverordnung: Kommentar* (CH Beck 2017) Art 15 Rn 4.
[56] ibid Art 15 Rn 30-31.
[57] Brent Mittelstadt, 'From Individual to Group Privacy in Big Data Analytics' [2017] Philosophy & Technology <http://link.springer.com/10.1007/s13347-017-0253-7> accessed 3 July 2017; Alessandro Mantelero, 'From Group Privacy to Collective Privacy: Towards a New Dimension of Privacy and Data Protection in the Big Data Era', *Group Privacy* (Springer 2017).

to relevant information about the scope and logic of automated processing. New approaches on how to protect 'group privacy'[58] in parallel to individual privacy need to be developed.

4.2 Second step: Transparent procedures to mitigate risks of profiling

5. Implement anti-discrimination tools and procedures

One of the most pressing problems concerning the IoT is discrimination.[59] Recital 39 GDPR reflects this concern as it states that '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' can lead to identification and profiling. This is further acknowledged in Recital 71 where it is stated 'the controller should use appropriate mathematical or statistical procedures for the profiling' to prevent discrimination or biases in profiling or automated decision-making. Sensitive or proxy data[60] as well as inaccurate or incomplete data[61] can form the basis for discriminatory effects,[62] especially when data sets are linked.[63] This is particularly challenging when a device has multiple users, as the usage behaviour of one user can inadvertently influence predictions about another user.

Critical assessment of the provenance of data is required. Consumers might provide incorrect data or do not fully understand the consequences if their behaviour is constantly monitored. Even when users are aware of the potential consequences of their usage of a device or service, changing settings may prove inconvenient or damaging.[64] Organisational measures should thus be implemented to guarantee the accuracy and reliability of the gathered data, while still ultimately deferring to the right of users to withhold private information (e.g. confirming whether or not a record is accurate) or knowingly provide false information for the sake of opacity or obscurity.[65]

---

[58] Mittelstadt and others (n 54); Brent Mittelstadt, 'From Individual to Group Privacy in Big Data Analytics' [2017] Philosophy & Technology; Alessandro Mantelero, 'Personal Data for Decisional Purposes in the Age of Analytics: From an Individual to a Collective Dimension of Data Protection' (2016) 32 Computer Law & Security Review 238; Bygrave (n 24) ch 15.
[59] Peppet (n 7) 117 ff.
[60] Barocas and Selbst (n 9).
[61] Moerel and Prins (n 28).
[62] Mittelstadt and others (n 54).
[63] Ruggieri, Pedreschi and Turini (n 31); Zarsky (n 31).
[64] Alessandro Acquisti and Jens Grossklags, 'Privacy and Rationality in Individual Decision Making' (2005) 3 IEEE Security & Privacy 26, 29–32.
[65] SA Bagüés and others, 'Disappearing for a While - Using White Lies in Pervasive Computing', *Proceedings of the 2007 ACM workshop on Privacy in electronic society* (2007) <http://www.scopus.com/inward/record.url?eid=2-s2.0-74049138834&partnerID=40&md5=85eabfb98cfe5d40f843c493ea3d460a>; Tene and Polonetsky (n 6); Evan Selinger and Woodrow Hartzog, 'Obscurity and Privacy' (Social Science Research Network 2014) SSRN Scholarly Paper ID 2439866 <http://papers.ssrn.com/abstract=2439866> accessed 18 February 2015.

Further, data processing can also lead to unexpected biases because potential relationships between data categories, often revealed only through aggregation and linkage of disparate datasets, may not be known at the time of data collection.[66] Tools such as ethical algorithmic auditing should be implemented to flag up discrimination.[67] Internal auditing schemes should be considered to guard against discrimination of protected groups, but also to protect victims of unanticipated discrimination.[68]

6. Tell users about assumptions and inferences

According to Recital 63 the right of access (Article 15) aims to make users 'aware of, and verify, the lawfulness of the processing'. Direct access to the data that is held should be given if possible. This allows not only for the accuracy of the collected data to be verified, but also rectified if inaccurate. In addition, Article 15(1)(h) allows data subjects to receive "meaningful information about the logic involved, as well as the significance and the envisaged consequences" of automated processing including profiling. Disclosing detailed information about the algorithms used for such processes could have adverse effects on data controllers' commercial interests, including trade secrets and IP rights.

Tools to provide users with meaningful information about the scope of data being processed and inferences being drawn from it should be implemented.[69] Existing mechanisms such as Google's ad manager[70] provide a starting point. Such tools should provide more than a general overview of profiling or automated decision-making supported by assumptions and inferences based on IoT data as is legally required. Rather, to help data subjects understand and manage the types of inferences being drawn, and to provide an opportunity to give additional information to correct inaccurate or undesired inferences, individual-level disclosures are preferable. This higher ethical standard for disclosing the process and results of inferential analytics can help optimise services (for users that choose to correct problematic inferences)

---

[66] Matt J Kusner and others, 'Counterfactual Fairness' [2017] arXiv:1703.06856 [cs, stat] <http://arxiv.org/abs/1703.06856> accessed 8 July 2017.
[67] Wachter, Mittelstadt and Floridi (n 20).
[68] Mittelstadt and others (n 54); Brent Mittelstadt, 'Auditing for Transparency in Content Personalization Systems' (2016) 10 International Journal of Communication 12.
[69] Wachter and Mittelstadt (n 9).Hildebrandt and Koops use the term "smart transparency." See: Mireille Hildebrandt and Bert-Jaap Koops, 'The Challenges of Ambient Law and Legal Protection in the Profiling Era' (2010) 73 The Modern Law Review 428.
[70] Google, 'Control the information Google uses to show you ads' [2017] Available at: <https://www.google.com/settings/u/0/ads/authenticated>.

and increase users' trust by rendering the most uncertain and unpredictable uses of IoT data more transparent.

7. Ethical data sharing practices

Privacy concerns do not necessarily rest solely with the data controllers that have initially collected the user's personal data via IoT devices and services. Rather, third parties with whom data controllers share the collected data can also pose a risk to privacy of users. As Weber explains, "since the possibility to build extensive personal profiles can be hardly avoided, data anonymization is important in the context of data sharing."[71] Insurance companies or employers[72] could, for example, have increasing interest to obtain data to assess current behaviour and infer future risks, for instance future likelihood of health impairments inferred from FitBit data. The GDPR requires that the recipients of data have to be disclosed if data sharing is planned (Art 13 and 14).

However, it is recommended that prior to sharing, an assessment of possible risks should be undertaken. Possibilities of, for example, racial and economic discrimination should be evaluated prior to sharing. Users might not foresee the possible risks of inferences drawn from their data, especially when datasets are shared with third parties and combined for related but distinct processing purposes. Edwards et al even suggest that "social impact assessment" should be considered that would "consider the public interest as well as the interests and rights of enterprises and users"[73] and look at factors like sharing practices since "B2B relationships, are not designed with privacy as a prime consideration."[74] Such assessments can help maintain privacy by opacity, ensuring that private information is not generated or shared with third parties beyond the terms agreed to by the user. Emergent risks, such as discrimination by proxy attributes or invasive inferences regarding unobserved aspects of the user's private life, should be identified and mitigated through such assessments.

8. Agile consent

Art 7 will readjust the power dynamic between data subjects and data controllers. Due to Art 7(4), the freedom with which consent is given will be evaluated on the basis of whether the obligation to share data is a precondition to use the service. This is in line with Art 13(2)(e),

[71] Rolf H Weber, 'Internet of Things: Privacy Issues Revisited' (2015) 31 Computer Law & Security Review 618.
[72] Peppet (n 7).
[73] L Edwards, D McAuley and L Diver, 'From Privacy Impact Assessment to Social Impact Assessment', *2016 IEEE Security and Privacy Workshops (SPW)* (2016) 56.
[74] ibid 55.

which requires data controller to state whether there is a 'statutory or contractual requirement, or a requirement necessary to enter into a contract, as well as whether the data subject is obliged to provide the personal data and of the possible consequences of failure to provide such data' in cases where data is collected from the data subject. In other words, privacy policies that deny the use of a service because a data subject refused to share all their data (e.g. pre-ticked boxes)[75] will no longer be legal. Such arrangements are also ethically questionable, as they undermine the user's ability to withhold private information.

In order to meet this requirement an agile and customised consent system is preferable.[76] It should be stated what kind of data is necessary for the offered service and what kind of data can be voluntarily shared. With this information in hand, users can make an informed choice about whether to personally use IoT offerings, or similarly to avoid IoT-enabled public spaces. For the latter, even if such disclosures are required, users will still be forced into a binary 'take it or leave it' choice unless specific forms of data collection can be disabled on a user-by-user basis. Of course, users must first be alerted to the presence of monitoring and the scope of data collection before being able to consent to specific aspects of monitoring, or avoiding the space altogether. Public spaces thus represent a unique challenge for balancing transparency, privacy via opacity, and the benefits of IoT devices and services.

9. Disconnecting and objecting

IoT devices are constantly collecting data about their users, which is why it has been suggested that disconnect options should be considered[77] that disable tracking. This approach is related to a similar provision in the GDPR: the right to object to profiling in Art 21. The framework states that regarding direct marketing purposes, the objection of a data subject will always trump the interest of data controllers. However, since profiling can be used for other purposes as well, e.g. to optimise services, data controllers can overrule the objection by demonstrating legitimate interests.[78]

However, data controllers are recommended to evaluate if profiling is necessary for their service and should possibly act according to the request of the user or at least consider

[75] de Hert and Papakonstantinou (n 29).
[76] Jane Kaye and others, 'Dynamic Consent: A Patient Interface for Twenty-First Century Research Networks' (2015) 23 European Journal of Human Genetics 141.
[77] Rolf H Weber and Romana Weber, *Internet of Things: Legal Perspectives*, vol 49 (Springer Science & Business Media 2010).
[78] Article 29 Data Protection Working Party, 'Opinion 06/2014 on the Notion of Legitimate Interests of the Data Controller under Article 7 of Directive 95/46/EC' (n 23).

opt-out options for specific purposes.[79] To incentivise data subjects to share and consent to processing of their data, data controllers should inform them about social or individual benefits, and provide options to forego these benefits by disabling data collection that is non-essential to the device or service's functioning. Doing so allows users to reasonably evaluate the opportunity costs of foregoing data collection for the sake of privacy.

4.3 Third step: Transparent contingency plans in case systems are compromised

10. Privacy by default and privacy by design

The GDPR states in its principles and various Articles (e.g. 6, 24, 32-34) that 'privacy by default' and 'privacy by design', pseudonymisation, encryption and other privacy enhancing tools (PET) should be used. This approach should help to increase user trust and public acceptance in identifying technologies, and enhance user privacy. However, as Ohm[80] stated, PETs are in most cases flawed. It must be assumed that a sufficiently motivated adversary will always be able to re-identify a user.

Rather than promising that personal data can always be protected, communicating realistic expectations of the extent to which their data can actually be protected should be considered. It is encouraged to explain that data protection will be guaranteed to the best of data controller's abilities. However, data controllers should explain that privacy risks will remain even under optimal conditions, thus providing users with a realistic assessment of whether their private information can actually be kept hidden. Explaining contingency plans in cases of data breach can help. For example, what measures are in place if systems are attacked? How will the negative consequences of data leaks be mitigated? It is also crucial to state how effective PET's will be in cases of cyber-attacks or data leaks. This type of transparency will help users to make an informed choice in deciding to use a service, as they will have more realistic expectations of the associated risks and mitigating factors. Similarly, it may force data controllers into adopting more robust contingency plans that can act as a selling point to potential users. If transparency and risk management enhance trust between data controllers and users, and greater trust leads to greater adoption of the technology, then detailed

[79] Alan Rubel and Kyle ML Jones, 'Student Privacy in Learning Analytics: An Information Ethics Perspective' (Social Science Research Network 2014) SSRN Scholarly Paper ID 2533704 <http://papers.ssrn.com/abstract=2533704> accessed 22 July 2015; Mireille Hildebrandt, 'Who Needs Stories If You Can Get the Data? ISPs in the Era of Big Number Crunching' (2011) 24 Philosophy & Technology 371.
[80] Ohm (n 10).

communication about how privacy by design and privacy by default have been implemented into a device or service can benefit both users and suppliers.

11. Be honest if (cyber)security fails

(Cyber)security hygiene is closely connected to protection of privacy. Security is one of the major problems in the IoT reflected in the principles and articles of the GDPR and the reviewed literature.[81] Article 33 GDPR will require data controllers to notify a supervisory authority when data breaches occur that are posing a 'risk to the rights and freedoms of natural persons'. However, data controllers only need to inform the data subject in serious cases where the consequences of the data breach will likely pose a 'high risk' to the data subject (Article 34).

Even though it is understandable that not every leak needs to be communicated, the barrier of 'high risk' should be seriously reconsidered, or at least granted a consistent operational definition. It remains unclear who will assess this risk, or how the consequences for users will be framed. Having a lower threshold for communicating data breaches could help to increase users' trust, otherwise they will not be aware of data breaches and leaks. IoT providers can develop internal definitions and codes of conduct to determine when 'high risks' exist, and what should be communicated to data subjects in those cases.

5 Two cases

To demonstrate how the guidelines described in the previous section could apply in practice, the rapid emergence of the IoT is analysed in two sectors: IoT used in public spaces as part of 'smart cities' initiatives, and connected cars as IoT devices. While the IoT is being applied in a myriad of sectors, these cases were chosen to show how application of the guidelines differs according to the types of space (e.g. public, private, and mixed) and users monitored (e.g. unidentified, known, and incidental 'users').

Smart cities present an additional challenge owing to the interconnectivity of multiple devices used to monitor public spaces and services, and potentially owned by public as well as private entities. 'Users' of smart city devices and services consist of individuals traversing a monitored area or engaging with a device (e.g. a smart rubbish bin), and can thus be conceived of as passive, sometimes unidentified individuals. Meaningful notification of the scope and purpose of data collection, as well as engagement with IoT devices or services themselves, is particularly challenging for devices embedded in an environment designed to passively and

[81] E.g. Weber and Weber (n 77); Ohm (n 10); Peppet (n 7).

unobtrusively monitor a place rather than specific, known users. In short, a meaningful choice of adoption is often impossible. The burden is placed upon 'users' to avoid monitored spaces if they do not wish to engage with the IoT. Compared with other 'smart' devices such as phones or connected cars, offering users a meaningful, non-binary mode of usage is particularly challenging for suppliers of smart city services. The same can be said of communicating potential risks and uses of the data collected, particularly when collected in as anonymous (rather than personal) data. IoT in smart cities is thus an interesting case to assess the proposed guidelines because of the particularly acute tension between provided a linked up, seamless service and offering meaningful transparency and data management to users to protect their privacy.

In contrast, connected cars often deal with known users and a well-defined set of sensors and potential uses of data. Obtrusiveness is less of a concern, as users either purchasing or hiring a connected car make an intentional choice to engage (assuming the sensing capabilities of the car are communicated beforehand). However, connected cars present their own challenges. While a single, well-defined 'bundle' of sensors, connected cars can nonetheless have multiple users, such as the members of a family or users of a particular car hire service. Known users can be joined by 'passive users', or other passengers that are potentially unaware of the sensing capabilities of the car and subsequent purposes and risks of data processing. A challenge is thus created for informational self-determination of the various active and passive cars of a car, as a single profile cannot be used for the car itself. Rather, differentiation between users is essential to ensure private information, such as usage records, is not inadvertently shared between users, and to enable users to exercise their individual data protection rights. A further challenge exists in resolving conflicts between the interests of passengers regarding their data protection rights, for example if one passenger requests deletion of activity data involving other passengers.

In both cases, an assumption is being made that the data collected by the IoT device or service is personal data. This assumption may prove false, particularly for IoT monitoring public spaces where obscuring identifying features at the point of collection may be preferable to the sparse options available to give 'users' an informed choice. For IoT that does not collect personal data, the legal provisions described above will not apply. They are thus considered to be outside the scope of this paper, and the guidelines proposed.

5.1 IoT in Connected Cities

Cities around the world are increasingly implementing IoT technologies.[82] Songdo[83] (South Korea), Masdar City (Abu Dhabi),[84] Barcelona,[85] London,[86] and Copenhagen[87] have all taken steps towards connected, data-driven services. Connected cities bear great societal potential for energy, water, transport, and waste management and assisted living with a potential global market of $408 billion by 2020.[88]

In Barcelona[89] for example, a combination of IoT devices and traditional open access datasets are being used to make city services 'smarter' and more efficient. WiFi is used by many smart transport services, including traffic management, real time timetables for buses, contactless payment, free parking spots, and rental of e-bikes. Smart waste management is enabled via sensors that signal when bins are full to help optimise logistics. Smart grid technology facilitates efficient energy generation, transmission, and distribution. In order to function, access is required to 'variety of devices such as, for instance, home appliances, surveillance cameras, monitoring sensors, actuators, displays, vehicles, and so on.'[90]

Due to the large-scale collection and sharing of data from various sources necessary for smart cities, privacy and security concerns arise.[91] Greater tracking of movements and activities of users is enabled by sensors monitoring public spaces, and nodes that allow users to connect their devices as they move across the city. As argued above, additional collection and sharing of personal data inherently creates additional risks to privacy. Large-scale privacy invasive assumptions can be drawn from smart city data without affected individuals being

[82] A Zanella and others, 'Internet of Things for Smart Cities' (2014) 1 IEEE Internet of Things Journal 22.
[83] 'Songdo IDB' <http://songdoibd.com/> accessed 8 January 2018.
[84] veolia.com, 'Masdar City, a Zero-Waste, Zero-Carbon City in the Desert' (*Living Circular*) <http://www.livingcircular.veolia.com/en/lifestyle/masdar-city-zero-waste-and-zero-carbon-desert> accessed 8 January 2018.
[85] Ross Tieman, 'Barcelona: Smart City Revolution in Progress' (*Financial Times*, 26 October 2017) <https://www.ft.com/content/6d2fe2a8-722c-11e7-93ff-99f383b09ff9> accessed 8 January 2018.
[86] Greater London Authority, 'Introduction to Smart London and Our Progress' (*London City Hall*, 8 March 2016) <https://www.london.gov.uk//what-we-do/business-and-economy/science-and-technology/smart-london/future-smart/introduction-smart> accessed 8 January 2018.
[87] Copenhagen Capacity, 'Smart City Copenhagen - a Living Lab' <http://www.copcap.com/set-up-a-business/key-sectors/smart-city> accessed 8 January 2018.
[88] Ove Arup & Partners Ltd, 'The Smart City Market: Opportunities for the UK; BIS RESEARCH PAPER NO. 136' <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/249423/bis-13-1217-smart-city-market-opportunties-uk.pdf> accessed 8 January 2018.
[89] Somayya Madakam and Ramaswamy Ramachandran, 'Barcelona Smart City: The Heaven on Earth (Internet of Things: Technological God)' (2015) 4 ZTE Communications 003.
[90] Zanella and others (n 82).
[91] For an in-depth overview of potential risks see Privacy International, 'Smart Cities: Utopian Vision, Dystopian Reality' (2017) <https://privacyinternational.org/sites/default/files/2017-12/Smart%20Cities-Utopian%20Vision%2C%20Dystopian%20Reality.pdf> accessed 13 February 2018.

aware, e.g. smart bins and eating behaviour; smart metering and sleeping schedules; and smart payment for transport and movement patterns. While smart cities can have significant societal benefits and efficiency gains, data subjects should be made aware of the potential privacy risks posed by greater collection of behavioural data, tracking and profiling of citizens.

Following the guidelines proposed here, several recommendations can be made to address the potential privacy risks of smart cities. By design, IoT devices that enable smart cities will collect and share data about large numbers of individuals who may be unaware that they are interacting with the IoT or generating data. DPIAs should be publicly accessible to ensure concerned individuals can subjectively evaluate the risks posed. As recommended in guideline 1, a suitable and well-known webpage should be chosen to reach a broad audience.

Following guideline 4, data subjects should also be granted individual-level control mechanisms as safeguards against profiling and tracking. Data subjects can exercise their right of access to learn about the data being collected, the purposes for which it is collected, and the logic involved in any automated decision-making. However, access requests could be technically challenging and potentially infringe the privacy of others. For example, data from shared waste bins can provide insights into the consumption habits of all tenants of a house.

Smart cities are therefore a good use case for the IoT to take group privacy (guideline 4) into account, in addition to individual privacy. The privacy interests of individuals and groups may, however, come into conflict. One approach to navigate competing privacy interests would be to inform individuals about the types of inferences that can or are intended to be drawn from their data, and the data of similar individuals (e.g. smart metering and people living in one household, sleep patterns; guideline 6). This could be done using 'counterfactual explanations' that can explain the influence of key variables on inferences made about the individual.[92] Protection should, however, also be afforded to groups. Inferential analytics can reveal personal aspects of groups or 'types' of people as opposed to individuals, and can lead to discriminatory effects on groups lacking formal legal protection under anti-discrimination law (e.g. all members of a specific area, all customers using public transport).

Guidelines 2 and 3 focus on how to communicate the risks of profiling and tracking. Usually these can be communicated via webpages and privacy notices. However, in the case of smart cities these forms of disclosure are often not feasible. Data collection will mostly

[92] Wachter, Mittelstadt and Russell (n 20).

happen passively and in public spaces without interactive notification. Despite this, data controllers must meet the notification duties imposed by the GDPR (Articles 13-14). As with CCTV, it may therefore be preferable to provide information about data collection and transmission via street signs using easily understandable icons with standard notices via QR codes in order to allow potential users to make an informed choice about whether to enter the monitored space.

As the challenges of notification demonstrate, smart cities also pose problems for agile consent (guideline 8) and the capacity of individuals opt-out (guideline 9). If data subjects opt out or refuse the give consent, the benefits of smart cities might be compromised. However, informational self-determination must be respected. Explaining the social benefits for users (e.g. less traffic congestion, increased health and well-being, lower utility bills) may help incentivise users to consent to the data collection and sharing necessary for smart cities to function. It is also unclear how agile consent and opting-out can function in public spaces in which entrants are monitored passively. Post-hoc opt-outs place an undue burden on users to locate suppliers and make an appropriate request, suggesting pre-emptive technological solutions or democratic decision-making regarding installation of IoT in public spaces may be more promising alternatives.

Trust between users and smart city developers and data controllers is also essential if users are expected to consent to data sharing.[93] Adhering to guidelines 5 and 7 can help increase trust. Data controllers should implement anti-discrimination tools (guideline 5) and give careful consideration to the parties and purposes for which users' personal data are shared (guideline 7). Smart cities collect and share data from diverse data sources, increasing the chance of privacy invasive insights in the life of individuals and groups. Location data enables sensitive inferences to be drawn about the socio-economic circumstances (e.g. restaurants visited) or ethnic background (e.g. postcode) of individuals. To prevent large-scale privacy invasive data analytics and discrimination, ethical auditing for algorithms will be essential to guarantee that automated systems are not biased (guideline 5).[94]

Moreover, this information should only be shared with trusted parties, and only if sufficient privacy by default and privacy by design standards (guideline 10) have been

[93] Unless one of the other grounds in Art 6 GDPR is used for lawful data processing.
[94] Mittelstadt, 'Auditing for Transparency in Content Personalization Systems' (n 68); Mittelstadt and others (n 54).

implemented. Smart cities often rely on WiFi and RFID,[95] which can be subject to cyber-attacks. In order to promote wide-spread deployment of IoT for smart cities, a robust infrastructure must be built (guideline 11) with open and honest communication if the system is compromised or hacked. Notification when systems are compromised is especially essential in smart cities due to the vast amounts of data collected from disparate data sources, as well as the potential harms for large parts of the population.

5.2 Connected Cars

Connected cars are a rapidly growing IoT sector. A recent report estimates that by 2020 connected cars will have a global market of EUR 115.26 billion with a strong focus on safety features, autonomous driving, and (personalised) entertainment (e.g. dash-boards) benefitting from Internet connectivity. Predicted benefits include higher safety standards (e.g. fatigue detection systems, tracking in case of theft), greater energy efficiency, increased productivity (e.g. spending less time en route), and improved convenience (e.g. remotely controlling thermostats).[96]

Despite these potential benefits, problems remain[97]. By definition, connected cars must be able to communicate with other cars, infrastructure, and other devices.[98] This connectivity introduces concerns around cybersecurity (hacking brakes, identity theft), behaviour monitoring as well as data protection and privacy.[99] These concerns could hamper widespread implementation, but the proposed guidelines can increase user trust and foster ethical business practices that protect user's privacy through transparency as well as personal opacity.[100]

Behavioural profiling (e.g. based on driving behaviour; or for entertainment services) is one major concern that can be mitigated through the guidelines. Guideline 1 proposes to make DPIA's public to allow users to inform themselves about the possible risks and assess whether data controllers have sufficiently addressed them. Guidelines 2 and 3 urge data

---

[95] Madakam and Ramachandran (n 89).

[96] Mark Lengton and others, 'Business Innovation Observatory- Internet of Things Connected - Cars' (2015) Case study 43 2–15 <https://ec.europa.eu/docsroom/documents/13394/attachments/2/translations/en/renditions/native>.

[97] Sandra Wachter, Brent Mittelstadt and Luciano Floridi, 'Transparent, Explainable, and Accountable AI for Robotics' (2017) 2 Science Robotics eaan6080.

[98] Lengton and others (n 96).

[99] HM Government, 'The Key Principles of Cyber Security for Connected and Automated Vehicles' (HM Government 2017) <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/661135/cyber-security-connected-automated-vehicles-key-principles.pdf> accessed 8 January 2018; Lengton and others (n 96).

[100] Privacy International, 'Is Your Car Connected?' (3 April 2017) <https://privacyinternational.org/blog/856./your-car-connected> accessed 13 February 2018.

controllers to use easily understandable language, but to elaborate when explaining possible risks, especially relating to profiling and tracking. This is crucial as data collection is seamless and ubiquitous. Data subjects may therefore be una/ware of the scope of evaluation undertaken, and unable to predict the inferences drawn about them. Data on braking and accelerating behaviour or location data can, for example, be used to create privacy invasive risk profiles of drivers.[101]

Guideline 5 recommends data controllers to implement anti-discrimination tools and procedures to mitigate unlawful and unintended discrimination to build trust between data subjects and data controllers. However, potential discrimination does not stop with the data controller that collected the data. Thus, guideline 7 is essential, as it calls upon data controllers to have ethical data sharing practices. This means that data controllers should not only inform data subjects about who they share the data with (as per Art 13-14), but also consider whether the data subject would find the recipients acceptable. For connected cars, driving behaviour data can have negative consequences for the data subject if shared with insurance companies to set personalised premiums or advertisers to serve tailored advertisements to in-car displays, for example.[102]

Data subjects should also be given the opportunity to exercise agile consent (guideline 8). In practice, data subjects would need to be able to customise the types of data being collected and processed by their car. But even after consent is given, data subjects should be given opportunities to opt-out, disconnect or disable tracking (guideline 9). Implementation of agile consent and opt-out features require data controllers to re-think their design choices. Ideally, devices and services can be designed to provide full or minimally limited functionality even if a data subject opts-out or does not give consent to the collection, processing, or sharing of certain data types. For connected cars, developers should consider giving drivers and passengers an option to disable the collection of location data, or to have multiple user accounts with customised preferences (e.g. if the car is used by a family rather than a single driver).

Guidelines 4 and 6 aim to grant data subjects oversight and control over how their data is evaluated. If data subjects exercise their right of access to learn about what and how data is being processed, data controllers should provide this information unless overriding interests

[101] Article 29 Data Protection Working Party, 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679' (2017) 17/EN WP 251 <http://www.hldataprotection.com/files/2017/10/20171013_wp251_enpdf.pdf> accessed 22 October 2017.
[102] On the possible discrimination due to advertisements see: ibid.

exist (e.g. privacy of others, trade secrets). These overriding interests should be interpreted narrowly.[103] However, exercising the right of access can be challenging as connected cars communicate with other cars, exchange and collect other users' data which may infringe on the privacy of others to which the individual may not be entitled. Where access cannot be granted, guideline 6 proposes to, at a minimum, inform individuals of the logic involved in data processing to explain what kind of inferences are being drawn.

Apart from privacy and discrimination issues, (cyber)-security is another significant area of concern that poses a barrier to trust in connected cars. Known vulnerabilities include remote hacking of brakes, 'virtual keys', and data theft (e.g. credit card details).[104] Here guideline 10 builds on the principle of privacy by default and privacy by design, advising data controllers to communicate how and to what extent privacy enhancing technologies help to protect privacy even if the system is compromised. This notion of open and honest communication is closely connected to guideline 11, which recommends that data controllers inform data subjects if the system is hacked or data leaks occurred, even if a high risk for data subjects is not expected. User awareness of the existence and effectiveness of security standards is essential to establish trust in safety critical systems such as connected cars.

6 Conclusion

IoT identification technologies raise many concerns around privacy and data protection. IoT systems relay on large data collection form diverse sources and data exchange with various devices to provide seamless, linked-up and personalised services. Machine learning and profiling is increasingly used to provide these personalised services. Pervasive data collection and linkage of disparate datasets enables invasive and unpredictable inferences to be drawn about individuals or groups. The inherent vulnerabilities of many IoT systems (due to limited processing power or a lack of commitment by developers) make them vulnerable to cyber-attacks.

The GDPR can help to alleviate many of the privacy risks posed by the IoT. The governing principles (Article 5) explain how European legislators envision fair, transparent and lawful data processing. As IoT systems can conflict with many of these principles (e.g. purpose and storage limitation, data minimisation), developers should be encouraged to go

[103] For an overview of legitimate interests of data controllers see: Article 29 Data Protection Working Party, 'Opinion 06/2014 on the Notion of Legitimate Interests of the Data Controller under Article 7 of Directive 95/46/EC' (n 23).
[104] Lengton and others (n 96).

above and beyond the strict legal requirements of the GDPR to design trustworthy and privacy enhancing systems and services that strives towards the GDPR's ideal of transparency. To assist developers in this process, a three-step transparency model in the form of eleven guidelines was proposed. Two case studies (smart cities and connected cars) were then analysed to show how the guidelines might apply in practice in contexts defined by qualitatively different spaces and types of users.

A fundamental tension exists between seamless and unobtrusiveness data collection and processing, and the privacy of users. As suggested in the proposed guidelines, transparency and meaningful disclosures by IoT suppliers and data controllers are essential to provide allow users to make well informed choices and manage their data effectively. Respect must be shown for users' right to privacy by ensuring they can withhold and withdraw personal data and forego usage of IoT devices and services as required. At the same time, a balance must be struck between disclosures in service of the right to privacy, and the privacy risks to data subjects incurred in pursuit of transparent operation. In other words, users' right to data protection and right to privacy must be balanced in the design and governance of identification technologies in the IoT.

IoT systems bear great potential in areas such as transport, health, energy consumption, public space and environmental monitoring, as well as personalised and linked-up services for users. In order to fully harness the potential of this technology, user trust and public acceptance is crucial. The proposed guidelines are a first step to dissolve some of the regulatory ambiguities about how to interpret the GDPR to protect user privacy without hampering the deployment of IoT systems. Going forward, IoT suppliers and data controllers must adopt consistent requirements and methods of disclosure, privacy by design and default, and risk assessments to engender trust with users minimise the trade-off between the benefits of the IoT and privacy. Ignoring this essential task risks imperilling the significant potential societal and economic benefits of the IoT. That users find these benefits worth the economic and privacy-oriented of pervasive monitoring cannot, and should not, be taken for granted.