THE UNIVERSITY
of EDINBURGH

Data Protection in the Age of Big Data

Legal Challenges and Responses in the Context of Online Behavioural Advertising

Jiahong Chen

PhD in Law – The University of Edinburgh – 2018

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where stated otherwise by reference or acknowledgment, the work presented is entirely my own.

Abstract

This thesis addresses the question of how data protection law should respond to the challenges arising from the ever-increasing prevalence of big data. The investigation is conducted with the case study of online behavioural advertising (OBA) and within the EU data protection legal framework, especially the General Data Protection Regulation (GDPR). It is argued that data protection law should respond to the big data challenges by leveraging the regulatory options that are either already in place in the current legal regime or potentially available to policymakers.

With the highly complex, powerful and opaque OBA network, in both technical and economic terms, the use of big data may pose fundamental threats to certain individualistic, collective or societal values. Despite a limited number of economic benefits such as free access to online services and the growth of the digital market, the latent risks of OBA call for an effective regulatory regime on big data.

While the EU's GDPR represents the latest and most comprehensive legal framework regulating the use of personal data, it has still fallen short on certain important aspects. The regulatory model characterised by individualised consent and the necessity test remains insufficient in fully protecting data subjects as autonomous persons, consumers and citizens in the context of OBA.

There is thus a pressing need for policymakers to review their regulatory toolbox in the light of the potential threats. On the one hand, it is necessary to reconsider the possibilities to blacklist or whitelist certain data uses with mechanisms that are either in place in the legal framework or can be introduced additionally. On the other hand, it is also necessary to realise the

full range of policy options that can be adopted to assist individuals in making informed decisions in the age of big data.

Lay Summary

This thesis addresses the question of how data protection law should respond to the challenges arising from the ever-increasing prevalence of big data. One form of personalised advertising, online behavioural advertising, is employed as an example of the application of big data. The research also reflects the latest data protection legal reform in the European Union. Big data may create both benefits and risks in personal, economic and political terms. While the current data protection legal system has been updated, it remains incapable of fully protecting the individualistic, collective and societal values potentially threatened by big data. Therefore, for policymakers to further improve the level of data protection, they should understand the regulatory options available within the current legal framework, as well as the additional measures that can be introduced.

Acknowledgements

Completing a PhD is a lonely but rewarding journey — this is what I have constantly been told by my senior colleagues, and now I cannot agree more. For me, it has been an exciting yet challenging adventure into a largely unexplored territory where only few people have set foot. Research decisions are to be made on a daily basis, often without a clear perspective where the chosen pathway would lead to. Yet, the sense of achievement is unrivalled as those key steps forward are made, or those little detours are completed.

Fortunately, throughout the journey I have received the most valuable guidance from some of the best mentors. I would like to say thank you — again — to my supervisors, Ms. Judith Rauhofer and Prof. Robin Williams, for all the most helpful support from them, both academically and personally. Gaining inspirations every time I have discussions with them has always been one of the greatest joys of my life as a researcher. I would also like to thank my thesis examiners Prof. Burkhard Schafer and Prof. Eleni Kosta, for their engaging, inspiring, encouraging and detailed discussion on my work. Needless to say, while every effort has been made to ensure the thesis is accurate and up-to-date, all remaining errors are solely my own.

Despite the sense of intellectual solitude — and sometimes frustration — as part of the research routine, the life as a PhD researcher has mostly been enjoyable. I feel extremely privileged to be surrounded by a group of excellent people from Edinburgh Law School. My postgraduate research directors, Dr. James Harrison, Dr. Andy Aydın-Aitchison and Dr. Filippo Fontanelli, have always been supportive and ready to help. I have had a lot of great memories with my colleagues from the PhD community, among whom my special thanks go to Himani Bhakuni, Israel Cedillo Lazcano, Shunyu Chi, Darin

Clearwater, Laurence Diver, Álvaro García Martínez, Pablo Grez Hidalgo, Emily Hancox, Matthew Jewell, Dawoon Jung, David Komuves, Wenlong Li, Lorna MacFarlane, Lucas Miotto Lopes, Jesus Manuel Niebla Zatarain, Fernando Pantoja Nunez, Daniela Rodríguez Gutiérrez, Xinxiang Shi, Qingxiang Wu, and Xiaoou Zheng.

During the two months of my visiting stay in Amsterdam, I received the warmest welcome from the members of IViR, University of Amsterdam. I would especially thank Prof. Nico van Eijk, Prof. Natali Helberger, Dr. Kristina Irion, Dr. João Pedro Quintais, Dr. Frederik Zuiderveen Borgesius, Tom Dobber, Max van Drunen, Sarah Eskens, Ronan Fahy, and Marijn Sax for their helpful feedback on my work.

I also owe my most heartfelt gratitude to those who, though not necessarily in physical proximity, have stayed closest to me and have offered the most important emotional support. Thank you, Ziyong Lin, Peng Xue, and Christopher Hack.

Last but certainly not least, I would like to say thank you to my parents, who have devoted their entire lives to loving, caring, supporting, and educating me. I love you.

Contents

Introduction

Context

Big data: Why does it matter?

In 2015, the leading research firm in the information technology sector, Gartner, decided to remove 'big data' from their Hype Cycle for Emerging Technologies,[1] a graphic representation designed to help businesses identify commercial opportunities alongside the adoption of technologies.[2] At the same time, machine learning made its first appearance in the Hype Cycle.[3]

Is big data a hype? Is big data *just* a hype? Is big data *still* a hype? Without explaining the exact meaning of big data (which we will get to in a moment), a quick overview of how 'big data' as a search term has been trending on Google in the last five years, compared against a few alternative terms, may give you a sense of the rise and fall of these technological trends. Since 2016, it seems that a number of new trends — such as machine learning, blockchain and Internet of Things (IoT) — have overtaken big data as the more popular search terms on Google.[4] If big data is just another bubble that is set to burst,[5] it would not make much sense to treat it as a unique phenomenon and

---

[1] Alex Woodie, 'Why Gartner Dropped Big Data Off the Hype Curve' *Datanami* (26 August 2015) <https://www.datanami.com/2015/08/26/why-gartner-dropped-big-data-off-the-hype-curve/> accessed 6 June 2018.

[2] Gartner, 'Gartner Hype Cycle' <http://www.gartner.com/technology/research/methodologies/hype-cycle.jsp> accessed 13 March 2018.

[3] Woodie (n 1).

[4] Google, 'Big Data, Machine Learning, Blockchain, IoT - Explore - Google Trends' <https://trends.google.co.uk/trends/explore?date=today%205-y&geo=US&q=Big%20Data,Machine%20Learning,Blockchain,IoT> accessed 6 June 2018.

[5] Mike Wheatley, 'Gartner Warns Big Data's Bubble May Burst as Enterprises Cut Investment' *SiliconANGLE* (4 October 2016) <https://siliconangle.com/blog/2016/10/04/gartner-warns-big-datas-bubble-may-burst-as-enterprises-plot-investments-elsewhere/> accessed 6 June 2018.

undertake a comprehensive study on it — not at least from a long-term policymaking point of view.

Yet, as will be shown throughout this study, the effects of big data are so profound that it has become part of many aspects of our daily life, and will probably continue to change even more aspects. While the exact phrase of 'big data' may eventually fade from the spotlight of public attention, the influence may, nevertheless, remain. The reason why big data is still something worth discussing in 2018, and something that should be taken seriously, may be that big data has been turning from something people talk about to something people experience. This explains why big data was dropped from the Hype Cycle. The decision, according to the analyst, did not mean that big data was no longer relevant, but rather that big data had 'become prevalent in our lives across many hype cycles.'[6] As big data is evolving and integrating with other emerging technologies,[7] its impacts on all parts of life will only become even more significant.

One of the latest impacts of big data that have captured public attention is the Facebook—Cambridge Analytica revelations. It was reported in late-2015 that user data collected from social media had been used by Cambridge Analytica for political campaigning during the US election.[8] This, however,

---

[6] Woodie (n 1).

[7] For the interactions between big data and such technologies as cloud computing, the internet of things, artificial intelligence and machine learning, see Divyakant Agrawal, Sudipto Das and Amr El Abbadi, 'Big Data and Cloud Computing: Current State and Future Opportunities' (Proceedings of the 14th International Conference on Extending Database Technology, Uppsala, 21-24 March 2011); Daniel E. O'Leary, '"Big Data", the "Internet of Things" and the "Internet of Signs"' (2013) 20(1) Intelligent Systems in Accounting, Finance and Management 53; Daniel E. O'Leary, 'Artificial Intelligence and Big Data' (2013) 28(2) IEEE Intelligent Systems 96; Justin Grimmer, 'We Are All Social Scientists Now: How Big Data, Machine Learning, and Causal Inference Work Together' (2015) 48(1) PS: Political Science & Politics 80.

[8] Harry Davies, 'Ted Cruz Using Firm That Harvested Data on Millions of Unwitting Facebook Users' *The Guardian* (11 December 2015) <https://www.theguardian.com/us-news/

did not raise a lot of public eyebrows, and still not so when another report in early-2017 further revealed the connection between Cambridge Analytica and the Leave campaign in the Brexit referendum.[9] Only when further media coverage followed in March 2018[10] did these practices eventually come under the spotlight.[11] Irrespective of the actual efficiency of their political micro-targeting,[12] the fact that a large volume of behavioural data is readily available for sophisticated analytics shows, at least partly, that we are already in an age of big data.

2015/dec/11/senator-ted-cruz-president-campaign-facebook-user-data> accessed 10 June 2018.

[9] Hannes Grassegger and Mikael Krogerus, 'The Data That Turned the World Upside Down' *Vice* (28 January 2017) <https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win> accessed 20 November 2017.

[10] Matthew Rosenberg, Nicholas Confessore and Carole Cadwalladr, 'How Trump Consultants Exploited the Facebook Data of Millions' *The New York Times* (17 March 2018) <https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html> accessed 10 June 2018; Carole Cadwalladr and Emma Graham-Harrison, 'Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach' *The Guardian* (17 March 2018) <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election> accessed 10 June 2018.

[11] Again, this is evidenced by Google Trends: The search term 'Cambridge Analytica' saw its first surge in March 2018. See Google, 'Cambridge Analytica - Explore - Google Trends' <https://trends.google.co.uk/trends/explore?q=Cambridge%20Analytica> accessed 10 June 2018.

[12] For different views on this, see Olivia Goldhill, 'The Psychology Behind Cambridge Analytica Is Massively Overhyped' *Quartz* (29 March 2018) <https://qz.com/1240331/cambridge-analytica-psychology-the-science-isnt-that-good-at-manipulation/> accessed 10 June 2018; Jonathan Allen and Jason Abbruzzese, 'Cambridge Analytica's Effectiveness Called Into Question Despite Alleged Facebook Data Harvesting' *NBC* (20 March 2018) <https://www.nbcnews.com/politics/politics-news/cambridge-analytica-s-effectiveness-called-question-despite-alleged-facebook-data-n858256> accessed 10 June 2018; Jane Wakefield, 'Cambridge Analytica: Can Targeted Online Ads Really Change a Voter's Behaviour?' *BBC* (30 March 2018) <https://www.bbc.co.uk/news/technology-43489408> accessed 10 June 2018; Privacy International, 'Cambridge Analytica Explained: Data and Elections' *Medium* (13 April 2017) <https://medium.com/privacy-international/cambridge-analytica-explained-data-and-elections-6d4e06549491> accessed 10 June 2018; Matthew Hindman, 'How Cambridge Analytica's Facebook Targeting Model Really Worked – According to the Person Who Build It' *Independent* (13 April 2018) <https://www.independent.co.uk/life-style/gadgets-and-tech/how-cambridge-analytica-s-facebook-targeting-model-really-worked-according-to-the-person-who-built-a8289901.html> accessed 10 June 2018.

Disclosures of governmental and commercial uses of personal data in recent years — such as the Snowden revelations — have played a key role in enhancing public awareness of the potential risks arising from the practices enabled by big data.[13] Yet, the short-term and long-term effects of the unfettered uses of big data, particularly when it comes to personal data, are yet to be fully exposed. The threats of the widespread application of big data concerning individuals, different groups of people and the entire society are still subject to further theoretical and empirical investigation. Well-grounded findings on these matters are of significant importance and urgency as they constitute the precondition for timely, informed regulatory approaches to big data. Without a coherent theory about the big data risks, it would be impossible to answer such questions as whether and how big data can and should be regulated.

Data protection law: A helpful approach?

When it comes to regulation, law is one of the most common regulatory instruments, although there are of course alternative policy tools available to regulators.[14] The utilisation of information — not only personal data — is often subject to the control of law, in fact across various sectors of law, such as privacy and defamation law, intellectual property law, state secrecy and freedom of information law, telecommunications and cybersecurity law, competition and consumer protection law, etc.

Among these potential legal approaches to the challenges of big data, data protection law stands out as a key branch of law that holds particular promise

---

[13] Ellen Daniel, 'Five Years On, What Has Changed Since the Edward Snowden Scandal?' *Verdict* <https://www.verdict.co.uk/snowden-scandal-five-years-gdpr/> accessed 21 June 2018.

[14] See Lawrence Lessig, 'The Law of the Horse: What Cyberlaw Might Teach' (1999) 113 Harvard Law Review 501.

for a number of reasons. First, the way data protection law functions is mainly through regulating the use of (personal) data directly. As will be revealed throughout this study (especially Chapters 4 through 6), by imposing a set of restrictions on the processing of personal data, data protection law has in effect created a systematic regime by which it can be determined whether certain operations on data are permitted. Second, certain principles embedded in data protection law may have significant implications — facilitative or restrictive — for the development of big data. On the one hand, the principle of free flow of personal data, for instance, is considered a crucial condition for the flourishing of big data.[15] On the other, the principles of data minimisation and purpose limitation (both will be analysed in Chapter 4) are sometimes considered contrary to the idea of big data.[16] The interactions between these principles and big data are therefore enormously important for understanding how law, in general, regulates big data. Third, data protection law is oriented by an open-ended collection of objectives and values that would allow reflection upon a range of interests that are either mutually-supportive or in conflict. The inclusiveness of data protection law is evident from the wording of the legislation. In the EU, for example, the current data protection framework — the General Data Protection Regulation (GDPR)[17] — pledges to protect 'fundamental rights and freedoms of natural persons and *in particular*

[15] Commission, 'A Digital Single Market Strategy for Europe' (2015) Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions COM(2015) 192 final, 14.
[16] Viktor Mayer-Schönberger and Yann Padova, 'Regime Change? Enabling Big Data Through Europe's New Data Protection Regulation' (2016) XVII The Columbia Science & Technology Law Review 315, 325-330; Tal Z. Zarsky, 'Incompatible: The GDPR in the Age of Big Data' (2016) 47 Seton Hall Law Review 995.
[17] 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] OJ L119/1 ('GDPR').

their right to the protection of personal data'[18] by setting out 'rules relating to the protection of natural persons with regard to the processing of personal data'.[19] The pluralism of the values covered by data protection law,[20] as long as they are related to the use of personal data, provides a helpful forum to accommodate different strands of discussions surrounding big data in a policy-informing context.

It is for these reasons that data protection law is chosen as the regulatory approach to big data for this study. That is not to say, however, that data protection law can address all issues resulting from big data. As the subsequent chapters unfold, it will become clearer that data protection law (or even more generally, law) has its limitations manifest on quite a few levels. Yet, despite these limitations, data protection law represents no less of a necessary and helpful pathway to making sense of the ideas and practices of legal solutions to the challenges of big data. This is the case even more, if an inclusive view of data protection is taken beyond some traditional, limited understanding of such concepts as, say, 'privacy'. It will help reveal a data protection paradigm almost as multi-faceted as the big data issues they are supposed to address, at least in part.

The legal picture of the data protection regime in many jurisdictions, however, remains in a state of uncertainty to a considerable extent. Having recently come into effect in May 2018, the EU's GDPR, for example, marks the outcome of a comprehensive reform of data protection law, with a large number of improvements to, and clarifications of, the existing regime codified into the new legislation. Still, as will be shown in the doctrinal analysis in Chapter 4, there is a degree of uncertainty about how the new law should be

[18] ibid art 1(2) (emphasis added).
[19] ibid art 1(1).
[20] ibid Recital 4.

interpreted or enforced in certain specific contexts. While the GDPR has largely retained the regulatory model created by its predecessor, the Data Protection Directive (DPD)[21], the sheer number of new provisions makes the model much more delicate. The ambiguities therefore come from both the old regime, where certain controversial legal issues remain unresolved, and also the new safeguards that give rise to additional challenges.

Moreover, and from a more critical point of view, another sense of uncertainty stems from the lack of a theoretical framework to evaluate the effectiveness of a data protection regime, especially in the face of big data. Such an absence results largely from, as highlighted above, the insufficient conceptualisation of the short-, medium- and long-term, individualistic, collective, and societal effects of big data with regard to the use of personal data. In order to have a full understanding of data protection as a legal approach to big data, the inquiry should therefore not stop at what the law *is*, but should also further ask what the law *can* and *should be*. Considering the profound implications of big data for private, communal and public life, the thorough examination of a data protection regime in the light of the big data risks will be of paramount importance.

Scope and expected contribution

Against the backdrop that big data has become an inevitable part of everyone's life with significant benefits and risks that are not always discernible, and that data protection law may be a useful regulatory approach to big data albeit perhaps still evolving, the following research question is of considerable value both theoretically and practically: *How should data protection law respond to the*

[21] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281/31 ('DPD').

challenges arising from the ever-increasing prevalence of big data? Of course, any attempt to address this whole challenging issue within one single study would prove overly ambitious and unrealistic. Indeed, both big data and data protection law are such broad topics that it would be impossible for one research project to cover both in full. Therefore, notwithstanding the extensive coverage that the research question might suggest above, the course of this study is narrowed down to specific social and legal contexts.

For big data, the case study of online behavioural advertising (OBA) is chosen as a representative of applied use of big data in the private sector. To put it simply, OBA is one form of online advertising that personalises the content of the adverts based on the user's behavioural patterns across websites. For example, after reading a post about a specific health issue on a website, an Internet user might later see an ad about a related healthcare product on a different website, as the advertising system concludes that, based on the user's browsing history, they might be interested in this product. The implementation of OBA is much more complicated in reality, as will be revealed in the next chapter, but the overarching idea is not difficult to understand: To optimise the performance of online advertising by making it tailored to the inferred interests of Internet users. The study of OBA provides a helpful perspective for the purpose of this study mainly for two reasons. For one thing, OBA represents a good case in point in that, as will be shown in the next chapter, it bears all of the features of a typical definition of big data. For another, the functioning of an OBA system demands high consumption of behavioural details, which allows data protection law to kick in as such details are arguably personal data. Both points will be further explained in detail (respectively in Chapter 1 and Chapter 4). It should also be noted that while OBA is largely considered a use case of big data in the private sector, the

concerns are not limited to private interests. It will be demonstrated throughout this study that even private uses of big data will have profound economic and political implications. This makes OBA an even better case study as it helps further reveal the multi-dimensional effects of big data without having to deal with the controversial and complex trade-offs in the cases of, for example, crime prevention or medical research.

On the legal side of the research question, the scope of data protection law is also limited to a specific context, or more precisely, a jurisdiction. The EU's GDPR is considered one of the most — if not *the* most — comprehensive data protection regime in the world.[22] The influence of EU data protection law outside Europe[23] is another reason why the GDPR can serve as a prototypical legal framework that could inform the ongoing debates around the world. Also, as mentioned above, the recent entry into force of the GDPR means that it is currently the most up-to-date outcome of policymaking on data protection issues, which presumably represents the latest equilibrium, if not compromise, of the interests of multiple stakeholders. Having said that, the choice of the GDPR as the main legal framework does not mean that it is taken for granted as the 'model' law for other jurisdictions. Rather, following the doctrinal analyses of the GDPR, the rest of the study will turn to an evaluative examination of this current regime.

With the more specifically-defined scope, the overarching research question can be rephrased as: How should data protection law, represented by the EU's General Data Protection Regulation, respond to the challenges

---

[22] Zarsky (n 16) 995; European Data Protection Supervisor, 'Data Protection' <https://edps.europa.eu/data-protection/data-protection_en> accessed 10 June 2018; Graham Greenleaf, 'Data Protection in a Globalised Network' in Ian Brown (ed), *Research Handbook on Governance of the Internet* (Edward Elgar 2013).

[23] See Graham Greenleaf, *Asian Data Privacy Laws: Trade and Human Rights Perspectives* (Oxford University Press 2014) 30-33; Alex B. Makulilo (ed), *African Data Privacy Laws* (Springer 2016) 18-19.

arising from the ever-increasing prevalence of big data in the context of online behavioural advertising? The study is expected to fill in a number of gaps in the literature in both theoretical and practical terms: a) it offers a further techno-economic account for the functioning of a typical form of big data; b) it undertakes a critical survey of the common justifications of the use of big data; c) it further theorises the potential risks of big data in a more systematic manner; d) it outlines the latest data protection legal framework with a use case of big data; e) it puts forward a set of criteria for the assessment of whether a data protection regime may sufficiently address the big data issues; f) it explores the regulatory possibilities to improve the current data protection framework with respect to big data. These original contributions will be seen even more clearly when the research question is divided into six sub-questions in a structured inquiry as illustrated below.

Structure and methodology

In order to fully answer the main research question defined above, the investigation will be conducted within a six-part structure, with each part addressing a sub-question that will contribute to the resolution of the overarching issue. The interdisciplinarity of this research means that a diversity of methodologies will be involved, and will vary notably from one chapter to another. It should be stressed from the outset that while this study has a strong empirical or socio-legal dimension, it does not involve first-hand collection of empirical evidence. Accordingly, some common research methods in social sciences, such as surveys, interviews or experiments, are not employed here. Yet, it does not follow that the enquiry is limited to traditional legal research method — the so-called 'black-letter' approach with the exclusive or at least major focus on interpreting the law with reference to legal sources like statutes or case-law. Instead, as will be seen throughout the study,

findings in empirical research are cited extensively (and of course, critically) to ensure the arguments are well-founded. References to non-scholarly sources (the 'grey literature') also form an important part of the research, including reports published by policymakers, regulators, businesses, trade groups and civil society organisations, as well as technical documents by platform developers and standards bodies. The topics and methodologies in the six main chapters are outlined in a roadmap as follows:

How has big data changed the digital realities? Chapter 1 sketches out an overall picture of the techno- and ecosystem of OBA as an instance of big data. It will begin by explaining the general concept of big data and the relevance of OBA as a case in point. A technological overview will be provided to show how a typical OBA system works in the tracking, profiling and targeting data lifecycle. This will be followed by further coverage of the value chain in the industry, explaining how certain critical players have gained enormous powers by growing in size, breadth and impact. The very specific technical or financial details about the OBA sector are not always accessible, but with the aid of existing academic works, as well as publicly-available information such as technical guidelines, financial statements and market reports, a lot can be revealed about the logics and impacts of the OBA sector.

How may big data potentially benefit individuals, the economy and the society? Based on the findings in the previous chapter, Chapter 2 critically examines the prevalent arguments in favour of the current practices commonly shared by OBA marketers. These arguments are usually made within a narrative of the potential benefits regarding the well-being of individuals, the growth of e-commerce, and the progression of the society, usually making reference to theoretical accounts as well as empirical evidence. The aim of this chapter is to review the reliability, relevance and transferability of these arguments in a

policy-informing context. Since the case for OBA, or big data in general, has been made both theoretically and empirically, by both researchers and stakeholders, the assessment of those findings will inevitably have to resort to a literature of an equally hybrid nature.

What are the latent risks of big data to individuals and the society? On the flip side of the discussions about the big data trade-offs, Chapter 3 conducts a thorough literature review on the theories of the perceived risks of big data. From an individualistic end of the spectrum to a societal one, a series of values will be revisited, ranging from privacy to informational self-determination, dignity, liberty, equality, power balance, and democracy. Again, the potential threats to these values will be exposed and further theorised with the case study of OBA. A metaphoric approach will be explored at the end of the chapter, with a view to a coherent conceptual framework to make sense of the genuine interests at stake.

How is data protection law regulating big data activities? Chapter 4 turns to the doctrinal analysis of the data protection legal framework in the context of OBA. As mentioned above, some key parts of the GDPR will be examined as a typical data protection regime. Following a brief historical and structural overview of the EU's data protection regulatory model, detailed analyses will be carried out on the applicability of the GDPR, the legal grounds for data processing, the data protection principles and the specific rules relevant to OBA. It will then be illustrated how the GDPR has inherited and strengthened the regulatory model of the DPD. Since online tracking serves as an important component of OBA and is subject to the regulation by the ePrivacy Directive,[24]

[24] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) (As amended by Directive 2009/136/EC) [2002] OJ L201/37 ('ePrivacy Directive').

a specific section will also cover the relevant part of that legislation. Throughout the chapter, the explanatory documents prepared by the EU legislative bodies, the case-law handed down by the Court of Justice of the EU (CJEU), the opinions and guidelines issued by the Article 29 Working Party (now the European Data Protection Board, EDPB), and the interpretation adopted by legal researchers will be cited as the main sources.

Can the current data protection law sufficiently address the risks of big data? With a relatively clear sense of the legal framework, Chapter 5 takes a more critical approach to evaluating the effectiveness of the GDPR in protecting personal, collective and public interests. In this chapter, individuals are viewed not just as autonomous persons, but also as consumers and citizens, whose autonomous, commercial and political interests may be compromised with the use of big data for OBA purposes. By associating different categories of interests with the diverse social roles individuals are playing, and assessing the law with regard to the level of protection afforded for these roles, it will become more straightforward to identify the areas where data protection law has fallen short of the expectations.

How can data protection law be improved to respond better to the big data challenges? The final chapter aims to bring together the findings in all previous chapters and shed further light on the possible way forward for the adaptation of the current data protection regime to the challenges of big data. A paradigm complementary to the existing one will be proposed in the hope of equipping regulators with better regulatory possibilities where the benefits and risks of big data can be better taken into account and the deficiencies of the GDPR can be at least partly remedied.

Chapter 1 How Big Is Big Data? The Techno-economic Landscape of Big Data in the Context of Online Behavioural Advertising

Big data is something old yet new. A biography of the idea of big data may begin with reference to 1944 when the escalation of library size came into notice.[1] Of course, the terminology of 'big data' had not been coined by that time, but the idea of what we call knowledge and information explosion today had been anticipated. The year of 2011 marks the beginning of the term's emergence in mainstream media. Quantifying public attention to big data by means of Google Trends statistics[2] would reveal how big data as a search term began to take the Web by storm from 2011, soaring all the way to its first peak in 2014 then remaining high-up to date. A parallel phenomenon has been taking place in academia as well. The number of publication indexed by Google Scholar mentioning 'big data' within 2010 was a mere 2,880, and 4,380 in 2011, then almost tripled to 11,900 in 2012, and then climbing all the way to 63,400 in 2017.[3] These figures are not necessarily the perfect indication of the public impact of big data, in particular considering how new technologies often experience the 'Hype Cycle' — a model developed by Gartner to illustrate the inflated expectations of emerging technologies.[4] The reality

[1] Gil Press, 'A Very Short History of Big Data' *Forbes* (9 May 2013) <https://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/> accessed 13 March 2018.
[2] Google, 'big data - Explore - Google Trends' <https://trends.google.com/trends/explore?date=all&q=big%20data> accessed 22 February 2017.
[3] Google, 'Google Scholar' <https://scholar.google.co.uk/scholar?q=%22big+data%22&as_ylo=2017&as_yhi=2017> accessed 13 March 2018.
[4] Gartner, 'Gartner Hype Cycle' <http://www.gartner.com/technology/research/methodologies/hype-cycle.jsp> accessed 13 March 2018.

might well be that the substance of 'big data' is overstated, simply because everyone seems to keep hearing and talking about it.

Yet, in the meantime, big data finds its way spreading to almost every sector. A wide range of industries are figuring out how much they can benefit from the potentials of big data, or at least trying to appear associated with big data so as to create an innovative image. This shows that the notion of big data might sometimes have been overused simply as a catchword. If the essential idea of big data had been foreseen in the 1940s, what it offers could turn out to be old wine in a new bottle, or a catch-all concept without much original substance. So why does 'big data' matter, in particular in non-technical terms? This chapter will address the following question: In what sense is big data 'big'? The enquiry will begin with a general understanding of big data, and then turn to a specific sector, namely online behavioural advertising (OBA). To illustrate how influential the industry of OBA — as a typical instance of big data — actually is, the investigation will be conducted both from a technological (Section 1.2) and an economic perspective (Section 1.3). A concluding section will summarise the state of affairs with particular emphasis on the mutually-supportive relations between the two dimensions, which forms a factual basis for the discussion about the potential benefits as well as risks of big data in the next chapters.

1.1 Big data and OBA: The big picture and the close-up lens

1.1.1 The rise of big data

As mentioned above, the concept of big data entered the public eye in around 2011 and then became a global trend in a matter of months. However, the taking shape of big data's key features is often accredited to a 2001 META Group (now Gartner) report, in which the analyst warns of the limits of

conventional data management and calls for industrial efforts to turn the challenges into opportunities.[5] Three crucial points are underlined in the report: the volume, velocity and variety of data ('the 3 V's'). These are not necessarily the defining characteristics of 'big data' — indeed, the term 'big data' did not appear in that report. One may easily point towards dozens of distinct versions of the big data concept,[6] but the emphasis on the 3 V's is relevant in most versions. The exponential development in all three dimensions of data use leads not just to a change in quantity, but a change in nature that is different from any previous phenomena in human history.

The first V, volume of data, as the most straightforward aspect, has remarkably increased over the past few years. It is estimated that in 1992, global internet traffic was 100GB per day, and that number remained the same in 2002, but *per second* this time.[7] In 2015, the figure was approximately 20,235 GB per second, and projected to triple to 61,386 in 2020.[8] That means by 2020, 'the gigabyte equivalent of all movies ever made will cross the global Internet every 2 minutes.'[9] Massive amounts of data is being generated, transferred, stored, shared, analysed and reused literally every millisecond. The large scale of data use can make a significant difference when it comes to extracting valuable information. It is one thing to have an idea of how a sample group of

---

[5] Doug Laney, *3D Data Management: Controlling Data Volume, Velocity and Variety* (2001) <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> accessed 20 February 2015.

[6] A summary of 12 of such definitions can be found at Gil Press, '12 Big Data Definitions: What's Yours?' *Forbes* (3 September 2014) <http://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours> accessed 22 February 2017. An interview with 43 industry leaders with regard to their understanding of the concept of big data can be found at Jennifer Dutcher, 'What Is Big Data?' (2014) <https://datascience.berkeley.edu/what-is-big-data> accessed 22 February 2017.

[7] Cisco, *The Zettabyte Era: Trends and Analysis* (2016) <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.pdf> accessed 11 August 2016.

[8] ibid.

[9] ibid.

individuals behave over particular periods, and it is another thing to possess the data of how *all* the individuals behave *all* the time. In fact, one of the most revolutionary philosophies of big data is *N=all*, signifying the aim to maximise data size, and eventually eliminating the very idea of sampling. Big data does not even have to be perfectly full or well-structured to be powerful; once it reaches a size big enough, useful patterns or correlations will surface. For instance, connecting clinical and cost data on a large scale helped prove the adverse effects of an approved drug in the US, which had taken years with small-scale studies when more than 27,000 heart attacks and deaths could have been avoided.[10] If combined with its application in other areas, big data is estimated to be able to create a value of $300 billion a year to the US healthcare system.[11]

Also important is the speedy feature, or velocity, of big data. With the help of big data, which is driven by the augmented processing power in collating, analysing and linking data at speed, it is now possible to extract valuable information from the massive data traffic in an almost instantaneous manner and put it into use in real time. Each link of the chain of data use — from raw data collection to final decision-making — takes time, and as data processing becomes increasingly intense, all the delays might add up significantly. Speed can be the lifeline for certain industries. Insurance companies, for example, may gain a decisive competitive edge by offering insurance quotes responsively.[12] How long a company keeps its potential customers waiting can make a huge difference. Marketing is another battlefield where timing is

---

[10] McKinsey & Company, *Big Data: The Next Frontier for Innovation, Competition, and Productivity* (2011) 41 <http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation> accessed 13 February 2017.
[11] ibid 49-51.
[12] PwC, *Insurance 2020: Turning Change into Opportunity* (2012) <http://www.pwc.com/gx/en/insurance/pdf/insurance-2020-turning-change-into-opportunity.pdf> accessed 22 February 2017.

crucial. At a time of a fast-paced lifestyles and constant change, any delay in detecting what comes under the spotlight could mean the loss of lucrative opportunities. A trending topic on the web may last only for a few hours, but if a start-up can seize the chance by, say, presenting its brand to interested individuals as the related keywords hit the headline — something Facebook has helped brands to achieve[13] — the benefits could be considerable. Once the delay of data processing is minimised to a particular point, big data will unlock a good number of business models that would be impossible if not implemented on a real-time basis.

Big data is also characterised by the variety of data. The increasing variety of data means that data is collected, verified, cross-referenced and synthesised from a much wider range of sources. The Avon Longitudinal Study of Parents and Children (ALSPAC, also known as Children of the 90s) represents a case in point. Starting in 1991 with over 14,000 pregnant women as participants, the project has been collecting and updating healthcare data for a generation.[14] Now they are taking one step further by linking their records to a wider range of administrative data from educational records to criminal convictions.[15] It would be natural to assume that, the more detailed and inclusive these profiles are, the more likely unexpected and precise findings can be discovered. Recently, based on ALSPAC data, nutritionists have confirmed the

[13] Garett Sloane, 'Facebook Lends Trending Hand to Brands' *Adweek* (17 January 2014) <http://www.adweek.com/digital/facebook-lends-trending-hand-brands-155060> accessed 22 February 2017.
[14] University of Bristol, 'Researchers' <http://www.bristol.ac.uk/alspac/researchers> accessed 13 March 2018.
[15] University of Bristol, 'Linkage to Routine Health and Social Records' <http://www.bristol.ac.uk/alspac/researchers/our-data/linkage/> accessed 13 March 2018.

correlations between academic achievement and breastfeeding,[16] something probably unanticipated at the time of launch of the project.

Datasets are now becoming increasingly inclusive, feeding on data from both public and private sectors, and both online and offline data points. The diversity of data sources is, to a greater extent, facilitated by technological as well as institutional developments. On the one hand, new technologies have enabled new channels for data collection (from PCs to mobile devices, from smart TVs to wearable gadgets), powerful infrastructure of data storage, and efficient mechanisms of data interoperability (such as protocols designed to standardise data exchange between data holders). On the other hand, governmental and commercial initiatives are also incentivising cross-domain data sharing. For instance, the UK is now pushing forward a standardisation project for sharing banking data, which would 'allow different software applications to communicate with each other and exchange data directly, without the need for human input each time.'[17] Acquisition of data from various holders can be also achieved by mergers and restructuring of businesses, just like how Google acquired video streaming site YouTube and blogging service Blogger, and expanded its kingdom to email service Gmail, online file storage Google Drive, mobile operating system Android and so on. The growing variety of data will therefore be unlike anything else we have

[16] University of Bristol, 'Insomnia More Common in Teens Whose Mums Dad Postnatal Depression' (2016) <http://www.bristol.ac.uk/alspac/news/2016/depression-and-insomnia.html> accessed 13 March 2018.

[17] Open Data Institute, *Data Sharing and Open Data for Banks: A Report for HM Treasury and Cabinet Office* (2014) 15 <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/382273/141202_API_Report_FINAL.PDF> accessed 13 March 2018.

experienced: It is not just combining a number of datasets or cross-referencing various files; it is aimed at connecting everything.[18]

Of course, as mentioned above, the 3 V's is but one way to define 'big data'. In fact, many commentators have interpreted big data in alternative ways, some of which are quite sceptical of the concept itself. For example, Ohlhorst brings in a fourth V in his version of big data — veracity — to flag up the danger of massive amounts of data 'lead[ing] to statistical errors and misinterpretation of the collected information.'[19] Indeed, some scholars have warned over the overestimation of the completeness of big data and the underestimation of the importance of traditionally collected data.[20] Besides the flaws of the dataset itself, the ways the collected data is processed are also prone to fallacies. It is suggested that big data algorithms might suffer from human,[21] systematic[22] and social biases[23]. Moreover, some critics have shown concerns about big data conclusions being used out of context,[24] in particular considering the mix-up of correlation and causation.[25] Further risks of big data

[18] Dave Evans, *The Internet of Everything: How More Relevant and Valuable Connections Will Change the World* (2012) <http://www.lehigh.edu/~inengrit/dropbox/eac1113/Cisco_Internet-of-Everything.pdf> accessed 22 February 2017.
[19] Frank Ohlhorst, *Big Data Analytics: Turning Big Data into Big Money* (John Wiley & Sons, Inc. 2013) 3.
[20] David Lazer and others, 'The Parable of Google Flu: Traps in Big Data Analysis' (2014) 343 Science 1203, 1205.
[21] Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard University Press 2015) 35.
[22] Danah Boyd and Kate Crawford, 'Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon' (2012) 15(5) Information, Communication & Society 662, 666-668.
[23] See ch 3, s 3.C, which highlights the potential threats of big data strengthening existing biases against stigmatised groups and creating new biases against worse-off groups in the society.
[24] Boyd and Crawford (n 22) 670-671.
[25] Ira S. Rubinstein, 'Big Data: The End of Privacy or a New Beginning?' (2013) 3(2) International Data Privacy Law 74, 76; Kate Crawford and Jason Schultz, 'Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms' (2014) 55 Boston College Law Review 93, 108.

will be developed in more detail in Chapter 3 with OBA as case study of big data. Before that, we need to explain why OBA actually serves as a good representation of big data.

1.1.2 Application to the Internet: OBA as a new trend

While it becomes clear that big data is all about being big in volume, velocity and variety, how big it really is remains quite ambiguous at this point. It is of course important not to lose sight of the bigger picture with respect to how big data is changing the digital reality in many aspects. Indeed, the examples given in the previous section may provide a rough idea of the ways big data is breaking the conventional boundaries of data use in scattered areas. Yet, in order to make sense of the power of big data, what is also needed is a typical case study that can capture the significance in all the three dimensions *individually*, as well as the combined effects when they are brought *together*. With these considerations in mind, the industry of online behavioural advertising (OBA) has been chosen as the use case for this study.

The instance of OBA bears all three essential hallmarks of big data: big volume, velocity and variety. In terms of the size of data related to online activity tracking, there is unlikely to be any precise figure indicating the overall size of data being collected. However, considering how much data is being collected from a single Internet user, there would be little doubt that OBA deserves the big data title. Google alone, for example, saves *all* the voice searches (not just the recognised speech, but the entire audio file), search keywords, and geolocations of *every* user unless they opt out or request deletion.[26] One who supposes they are not using Google services might

[26] Becca Caddy, 'Google Tracks Everything You Do. Here's How to Delete It' *Wired* (14 August 2017) <http://www.wired.co.uk/article/google-history-search-tracking-data-how-to-delete> accessed 13 March 2018.

probably be still being tracked by Google-related products, such as Google Analytics.[27] It is estimated that tens of millions of websites are using Google Analytics,[28] which records users' visits and interactions on these sites for analytical and (partly) advertising purposes.[29] As will be shown below in the following sections, one single visit of a webpage may trigger multiple connections that carries a wide scope of information.

When it comes to the speed of data use, OBA also represents a case in point. Unlike what an Internet user might assume, the adverts on a webpage do not come up at the same time the page is loaded. In fact, as will be explained in Section 1.2, after the webpage finishes loading, dozens of additional communications will immediately take place between several actors within the advertising network so as to determine, say, what might interest the user, what the webpage is about, what product to present to the user, and what content on the advert. The speed at which such analyses and communications have been performed is so optimised that a user can hardly notice the delay.

Lastly, the big variety feature is also manifest in the case of OBA in that data collection is conducted across websites (whether news or travel website), across devices (whether PC or smartphone), across services (whether Google Maps or YouTube), across channels (whether provided by the users themselves or by a data agent) and across formats (whether structured data like visit records or unstructured data like email content). All the information from these sources will end up being fed into the generation of both individual profiles and general demographic models. As such, OBA matches all the major

---

[27] Google, 'How Google Uses Data When You Use Our Partners' Sites or Apps' <https://www.google.com/policies/privacy/partners/> accessed 13 March 2018.
[28] Matt McGee, 'As Google Analytics Turns 10, We Ask: How Many Websites Use It?' *Marketing Land* (12 November 2015) <https://marketingland.com/as-google-analytics-turns-10-we-ask-how-many-websites-use-it-151892> accessed 13 March 2018.
[29] Google, 'Policy Requirements for Google Analytics Advertising Features' (2016) <https://support.google.com/analytics/answer/2700409> accessed 13 March 2018.

defining features and thus may serve as a useful approach to understanding big data.

The study of OBA may provide beneficial insights into big data also because it does not only mirror the technically revolutionary aspect of big data, but the socio-economic drives as well. It is already pointed out above that big data results not just from technological developments. Governmental and non-governmental inputs also have a significant role in the rise of big data. In the case of OBA, the massive scale of data collection is largely enabled by the fact that a group of mega-players in the industry have secured sufficient powers to implement big data. These powers include the infrastructural resources, market penetration and standard-setting capability that are necessary for the widespread application of their big data agenda. To such an extent, OBA — and big data — should not be viewed only as a technological artefact, but rather a social engineering process that significantly influences individual behaviour and the distribution of social wealth and cultural capital. In the following two sections of this chapter, the landscape of the OBA industry, as an instance of big data, will be outlined from both a technical and an economic perspective. Section 1.2 will address the technical aspect of the day-to-day practices of a typical OBA system, explaining how data is generated, transferred and consumed between involved parties and eventually turned into revenues. Section 1.3 will then turn to a macro level, illustrating how key players have built up big data empires by virtue of market domination, business expansion and industrial influence.

1.2 The known and the unknown: Technological model of OBA behind the scenes of the Internet

The striking success that the leaders in the OBA industry have won owes much to the availability of new technologies. At the end of the day, any ground-

breaking business models would be subject to technical constraints. A fully informed image of the industrial power entails the understanding of how Internet technologies have both restricted and enabled OBA attempts, and how the former has become circumvented and reshaped by the latter. This section will serve as a very brief overview of the technical structure in plain language as far as possible. The technical and business models differ drastically from one ad network system to another, but most of the popular ones share a similar three-stage process: tracking, profiling and targeting.[30]

The illustration in this section relies on multiple sources: publicly available documents released by OBA network providers, technical documents (including released source codes) of the programme framework, patent records and secondary references. It is true that not all the technical details are visible to an external observer, in particular with the generic information designed for ordinary users. However, thanks to the standardisation of information exchange between the actors in an OBA network, a lot of information may still be distilled from such protocols (the 'interfaces' or 'APIs'). For instance, in the tracking stage, it is technically possible to monitor the data communications between the browser and remote servers. As regards targeting, given the necessity to automate the data connection between the OBA network provider and the advertisers (or their agents), the underlying portals must be clearly and openly stated to the latter, not necessarily in natural language, but at least in the form of codes. This opens up a window for us to look into the operative details. The profiling phase might be less

[30] Zuiderveen Borgesius divides the general practices into five phases: data collection, data storage, data analysis, data disclosure and targeting. See Frederik J. Zuiderveen Borgesius, *Improving Privacy Protection in the Area of Behavioural Targeting* (Kluwer Law International 2015) ch 2. The Article 29 Working Party explains the functioning of an OBA system with three main parts: distribution (ad delivery) system, tracking technologies and profile building. See Article 29 Data Protection Working Party, 'Opinion 2/2010 on online behavioural advertising' (2010) 00909/10/EN WP 171, 4-7.

transparent, as most of it is performed within the servers internally. The difficulties of achieving insights into the internal operation is exacerbated by the fact that most, if not all, OBA profilers consider their algorithms as trade secrets and therefore would hardly make them available to the public. Having said that, for the purpose of this research, there is no need to review the profiling mechanisms in full detail. An overall idea of how Internet users are being observed and ranked will suffice. Putting together information from indirect sources, such as patent files or secondary sources, a lot can still be reasonably revealed about the logics of such technologies.

1.2.1 Tracking

The most important precondition for an OBA system to operate is that users' online behaviour is trackable. However, this feature should not be taken for granted, considering the technical constraints at the infant stage of the Internet. The HTTP Protocol, which underpins the functioning of the vast majority of web browsing today, has been designed to be 'stateless'.[31] That means the original technical standards did not enable a server to maintain a dialogue with a specific user. When a website receives a webpage request (probably as result of a click on a link or an entry in the address bar), it cannot distinguish whether this is sent by a new user or a returning one that it served before. Under this condition, the 'log-in' or 'shopping cart' features were impossible.[32] A registered user who has submitted their username and password on a first request will become completely unrecognisable to the server on their next movement on that website. The server simply cannot tell a user (who has just

[31] D. Kristol and L. Montulli, *HTTP State Management Mechanism* (2000) <https://www.ietf.org/rfc/rfc2965.txt> accessed 13 March 2018.
[32] Federal Trade Commission, 'Protecting Consumer Privacy in an Era of Rapid Change: A Proposed Framework for Businesses and Policymakers (Preliminary FTC Staff Report)' (2010), 13.

given the correct login credentials) apart from others, since the architecture does not enable the server to identify the user.

To overcome this disadvantage, a patch was proposed in 1994 and adopted by mainstream web browser producers later.[33] The solution involves a short message sent from the server along with the content of a webpage and stored in the user's web browser. The idea is that when a new user visits a website, the server will designate a unique ID to this user and attach it to the main text of the webpage in a way invisible to the user ('HTTP header').[34] The browser then remembers that ID, and sends it back along with any subsequent requests to the same website.[35] Next time the server handles a request, it can authenticate the user by recognising that ID. This technique is named 'cookies'. With cookies, the activities of Internet users are trackable and recordable. The lifespan of a cookie varies depending on the expiry date specified by the server. It can be as short as expiring immediately the browser is closed ('session cookie') or as long as almost forever ('persistent cookie').[36] Once the cookie expires or is cleared by the user, usually the server will lose track of that user as it can no longer recognise the same user on their next visit.[37]

---

[33] Steven C. Bennett, 'Regulating Online Behavioral Advertising' (2011) 44 The John Marshall Law Review 899, 900.

[34] Kristol and Montulli (n 31) 4-5.

[35] ibid 8-9.

[36] Article 29 Data Protection Working Party, 'Opinion 04/2012 on Cookie Consent Exemption' (2012) 00879/12/EN WP 194, 4. An EU study finds that some cookies are set with a duration period of nearly 8000 years. See Article 29 Data Protection Working Party, 'Cookie sweep combined analysis' (2015) 14/EN WP 229, 2.

[37] Combined with 'flash-cookies' or other forms of device fingerprints, however, erased cookies can be 'respawned'. See Ashkan Soltani and others, 'Flash Cookies and Privacy' (2009) <https://ssrn.com/abstract=1446862> accessed 7 January 2017; Article 29 Data Protection Working Party, 'Opinion 2/2010 on online behavioural advertising' (n 30) 6-7; N. van Eijk and others, 'Online Tracking: Questioning the Power of Informed Consent' (2012) 14(5) Info 57, 14-15; Omer Tene and Jules Polonetsky, 'To Track or "Do Not Track": Advancing Transparency and Individual Control in Online Behavioural Advertising' (2012) 13(1) Minnesota Journal of Law, Science & Technology 281, 293-294.

For safety reasons, the cookies placed by a website are only accessible to the same website, or more precisely, only to webpages that share the same domain name with the one that sets the cookies. That means the tracking on one website is supposed to be completely separate from that on another website. However, a workaround had been developed to bypass such safeguards. The key to the circumvention of this restriction is to make references from the webpage of one site to another. For instance, when a user visits *food.com/pasta.htm*, that page may contain a hidden reference to a file of its advertising partner, say, *ads.com/tracking.htm*. When the browser tries to load the *tracking.htm* file, *ads.com* will have the opportunity to set a cookie on the user's device. If later the same user visits another website, say *movies.com/comedies.htm*, and *movies.com* also happens to work with *ads.com* by also including a hidden reference to *ads.com/tracking.htm*, then *ads.com* will be able to read the cookie that it has previously set. This way, *ads.com* may recognise the user and learn that they have viewed both *food.com/pasta.htm* and *movies.com/comedies.htm*. This type of cookie is called a 'third-party cookie' and is widely used for advertising or analytical purposes.[38]

It should be noted, however, that the cookies set respectively by *food.com* and *ads.com*, although done almost simultaneously in one loading of the webpage, are usually entirely separate and inaccessible to each other. Also, the embedded reference to *ads.com* can only be allowed on a page-by-page basis. That means, if the reference is made on one page, say, *food.com/pasta.htm* but not another *food.com/pizza.htm*, the user's visit to the latter will be undetectable and thus *ads.com* will not be able to tell that the user has also viewed a page of pizza. That means, not all the activities of a user on *food.com* are visible to

[38] Article 29 Data Protection Working Party, 'Opinion 2/2010 on online behavioural advertising' (n 30) 6.

ads.com; *food.com* may decide to exclude the tracking from sensitive pages, such as *food.com/reset_password.htm*. Despite these restrictions, there is still a massive amount of information that can be extracted from the browsing history[39]. Some data is very revealing, such as the URLs of pages that the user has viewed, how the user interacts with the elements on the page, the IP address of the user and so forth;[40] others can be even more intrusive, including even the cursor trajectory outside the web browser.[41]

By means of third-party cookies, cross-site tracking of Internet users is possible. As a user surfs from one site to another, their interactions with all these websites are all kept under tracking so long as they share the same OBA network or, as the case might be, subscribe to different networks that share data. What might make the situation even more disturbing is that major OBA operators also provide a multitude of services, which facilitate tracking to an even greater extent. Google, for instance, runs searching (Google Search), email (Gmail) and video (YouTube) services, and a user's activities on these affiliated sites may in theory — and actually in practice for some of these services — be placed under a same tracking mechanism. The combination of searching with the keyword 'fishing point', viewing of an article featuring fishing skills and subscribing to a fishing-lover mailing list can be a very strong indicator of potential interest in a fishing kit, which can be fed into the profiling stage. Moreover, as new forms of tracking techniques — such as

---

[39] Zachary Weinberg and others, 'I Still Know What You Visited Last Summer: Leaking Browsing History via User Interaction and Side Channel Attacks' (2011 IEEE Symposium on Security and Privacy, Berkeley, 22-25 May 2011).

[40] Google, 'Privacy Policy' (2017) <https://static.googleusercontent.com/media/www.google.co.uk/en/uk/intl/en_uk/policies/privacy/google_privacy_policy_en_uk.pdf> accessed 26 September 2017.

[41] Jon Gold, 'IE Exploit Can Track Mouse Cursor Movements - Even When You're Not in IE' *Techworld* (13 December 2012) <http://www.techworld.com/news/security/adobe-releases-security-updates-for-flash-player-coldfusion-3416178/> accessed 27 June 2017.

'device fingerprinting' that combines hardware or software specifications to identify a unique device [42] — are emerging as a complement to, or substitute of, cookies, tracking becomes even more ubiquitous. For example, Google has introduced the 'Advertising ID' feature to its mobile platform Android, so tracking can be also done within Google Play apps (where cookie does not work well).[43] Some trackers also use cookie-like functionalities provided by web browser plugins, such as Flash Player, to create the so-called 'Flash cookies', which fall outside of the browser's cookie control.[44]

1.2.2 Profiling

Businesses' endeavours to keep track of Internet users' online motions are not invested out of mere voyeurism. They are rather driven by commercial incentives. Maybe they are not particularly interested in the users' private life, but they are certainly hungry for information about what might interest their potential customers. That is why major ad network providers like Google and Facebook are 'interest-based',[45] focusing on discovering an individual user's consumption interests. To achieve greater precision of their predictions, they maintain a full list of all relevant business categories and manage the information of the users to see which of these categories are likely to match their interests.[46] Such a process involving the organisation of the data collected from users is termed 'profiling' in this study.

---

[42] See Article 29 Data Protection Working Party, 'Opinion 9/2014 on the application of Directive 2002/58/EC to device fingerprinting' (2014) 14/EN WP 224.

[43] Google, 'Advertising ID' <https://support.google.com/googleplay/android-developer/answer/6048248> accessed 13 March 2018.

[44] Tene and Polonetsky (n 37) 292-294.

[45] Google, 'About Personalised advertising' <https://support.google.com/adsense/answer/113771> accessed 13 March 2018; Facebook, 'What Is Online Interest-based Advertising from Facebook, and How Can I Control Whether I See Online Interest-based Ads?' <https://www.facebook.com/help/164968693837950> accessed 13 March 2018.

[46] Google, 'Topics Used for Personalised Ads' <https://support.google.com/ads/answer/2842480> accessed 13 March 2018; Google, 'How Does Facebook Decide Which Ads to Show

For the marketing industry, the most important key to the effective profiling of users is 'the seamless integration of consumer data across offline and online touch points.'[47] In their eyes, the whole point of building up massive profiles of their users lies in the need to digest such data and strategise their marketing approaches accordingly. From that point of view, fragmented data about users, no matter how big the volume can be, is entirely useless unless they can be aggregated in a way that can inform decision-making. The user identities are the nexus that connects the dots inside the ocean of online behavioural data. For the purpose of online marketing, it is not necessary to know the users' real names,[48] but it is necessary to figure out which activities are committed by the same user, and what inferences can be drawn from such activities.

How marketers make sense of the data they have collected is usually opaque to users or even their business partners. The algorithms behind the process of putting together massive analysis of data from various channels at speed are the key to optimising profiling performance, and thus the most important assets of a data-driven company. Also, the profilers may claim that the logics of these algorithms are too complicated for average people to understand,[49] which could be true given how fast machine-learning systems can grow. However, these arguments — namely trade secrets and complexity — do not fully justify the lack of transparency. As will be shown in Chapter 3,

---

Me and How Can I Control the Ads I See?' <https://www.facebook.com/help/562973647153813> accessed 13 March 2018.

[47] Craig Dempster and John Lee, *The Rise of the Platform Marketer: Performance Marketing with Google, Facebook, and Twitter, Plus the Latest High-Growth Digital Advertising Platforms* (John Wiley & Sons, Inc. 2015) 42.

[48] This however does not mean that they do not or cannot link a user's profile to their real name or other identities. See Omer Tene, 'What Google Knows: Privacy and Internet Search Engines' [2008](4) Utah Law Review 1433, 1448-1449.

[49] Danielle Keats Citron and Frank Pasquale, 'The Scored Society: Due Process for Automated Predictions' (2014) 89 Washington Law Review 1, 5; Pasquale (n 21) 15.

black-boxing data processing would pose serious threats to individual and public interests. In the offline world, a manufacturer is simply not allowed to refuse to provide information on its products that is vital to consumers on the ground that it would go against business interests or cause extra costs — which holds good for OBA as well. Despite such opacity, however, it is still possible to extract valuable information from publicly available materials on how the system works. For the purpose of this study, it will suffice to have a general conception of the way a typical interest-based OBA profiling system operates. This would probably not be precise enough for experts or the general public to carry out a thorough audit on the accountability of such systems, but that would be a different context.

A Google patent file has documented a probable technical solution for the profiling of Internet users.[50] In this solution, two core modules support the building of user profiles: User Interest and User Demographics.[51] The User Interest Module calculates and updates a given user's interest profile by accounting for every action recorded from the user. For example, the viewing of a sports-related webpage would result in an additional point to the 'sports' and 'entertainment' categories of the viewer's profile. The relevance of an action by the user to these categories is measured in a weighted manner. That means when deciding how much the action would affect the strength of an interest category, the type of the action (e.g. viewing, clicking, buying, etc.), the relevance of the page (e.g. more sports-related and less entertainment-related), the frequency of such actions in the given period and other factors

[50] Xuefu Wang and others, 'United States Patent: Generating User Profiles' (2013) <https://patentimages.storage.googleapis.com/87/fb/2c/bf0a5e5a68f605/US8352319.pdf> accessed 13 March 2018.
[51] ibid fig 1.

will be taken into account.[52] Over time, the module is able to predict how likely a particular user is interested in each of the categories.

The User Demographics module aims to figure out the demographic features of a given user. According to the patent document, such information includes 'geographic location, age, gender, income range, household income range, size of household, maximum educational attainment, children in the household, etc.' [53] These indicators are determined not just from the information given by the users, but also their online behaviour as well. This is achieved in a way similar to how the User Interest module works. Each of the user's activities will contribute to the inference that the user belongs to a particular demographic bracket.[54] Working together, the two modules may establish quite an accurate profile for every user. It should be noted that just because Google has been awarded this patent does not mean that this characterises exactly their solution currently in use. Also, there would hardly be any information about the weighting of individual elements as such details would certainly be kept as trade secrets. However, this model matches the less detailed information provided on their official website,[55] as well as other secondary sources.[56] Thus, it is reasonable to believe that this has largely represented the functioning of Google's system.

---

[52] ibid 5-6.

[53] ibid 8.

[54] ibid 8-9.

[55] Google, 'How Google Infers Interest and Demographic Categories' <https://support.google.com/adsense/answer/140378> accessed 13 March 2018; Google, 'About Google Ads' <https://support.google.com/ads/answer/1634057> accessed 13 March 2018; Google, 'About Ads Based on Websites That You've Visited' <https://support.google.com/ads/answer/1697735> accessed 13 March 2018.

[56] Claude Castelluccia, Mohamed-Ali Kaafar and Minh-Dung Tran, 'Betrayed by Your Ads! Reconstructing User Profiles from Targeted Ads' in Simone Fischer-Hübner and Matthew Wright (eds), *Privacy Enhancing Technologies* (Springer 2012) 3-8; Zuiderveen Borgesius (n 30) 41-44; Eric Siegel, *Predictive Analytics : The Power to Predict Who Will Click, Buy, Lie, or Die* (Wiley 2016) 30-36.

It has been highlighted in the discussion of tracking how ad network providers make use of 'third-party' cookies to carry out cross-site tracking. One particular technical restriction is that cookies that are set by one website (*food.com*) are not accessible to one another (*ads.com*). With this barrier, although both the ad publishers (*food.com* and *movies.com*) and ad network providers (*ads.com*) may keep separate records of the same user, they cannot merge these profiles into one, as the IDs are not compatible in these systems. However, this rule of separation is circumvented with just a few more technical tricks. For instance, when *food.com/pasta.htm* is making a reference to *ads.com* to allow the latter to set up its own cookies, it may refer to a URL of *ads.com/tracking.htm?foodie_id=13579*. The additional parameter 'foodie_id=13579' confirms to *ads.com* that this user (say, User No. 24680 in *ads.com*'s database) is actually the same person as User No. 13579 in *food.com*'s database. By doing this, *ads.com* can retrieve more data (such as demographic data) about that user from *food.com* with their ID 13579. Such a practice is known as 'cookie matching'.[57]

The use of cookie matching has profound implications for the entire online profiling ecosystem, not just in a technical sense, but in a financial sense as well. It creates the possibility of profiling on multiple levels and selling of cookies in an intertwined network. A wider group of profilers other than advertisers, ad publishers and ad network providers are brought into play. Suppose a user is viewing a recipe on *food.com*, who is a partner of a data provider[58] *data.com*. By means of cookie matching, *data.com* gains access to the

[57] Joseph Turow, *The Daily You: How the New Advertising Industry is Defining Your Identity and Your World* (Yale University Press 2012) 80-81; Mike Smith, *Targeted: How Technology Is Revolutionizing Advertising and the Way Companies Reach Consumers* (American Management Association 2015) 74-75.

[58] For an exemplary list of data providers, see Ramsey McGrory, 'The Data Providers: One Quadrant Chart To Rule Them All' *AdExchanger* (21 February 2013) <https://adexchanger.

user's cookies as well as further information from *food.com*, and in return, *food.com* gets paid.[59] By the same token, *data.com* allows a data aggregator[60] *aggregate.com* to collect such information, of course, with a charge. Advertising agent *agent.com* happens to be a customer of *aggregate.com* so it buys the access to this user's profile. This way, at least four profiles have been created — and sold — for at least four times,[61] respectively by *food.com*, *data.com*, *aggregate.com* and *agent.com*.

Perhaps the even more disturbing fact is that cookie matching enables these profilers to identify the same user in other profilers' databases. Let us assume, for instance, a data provider (*data.com*) has expertise in, say, lifestyle websites (including *food.com*) while another provider (*data2.com*) is focused more on the entertainment sector (such as *movies.com*). If they are both *aggregate.com*'s data sources, then the latter would be able to recognise the same user from both *data.com* and *data2.com*'s profiles, and thus create a fuller profile that covers this user's dietary as well as recreational preferences. The reality can be even more complex with a greater number of actors in the

com/data-driven-thinking/the-data-providers-one-quadrant-chart-to-rule-them-all> accessed 22 February 2017.

[59] Data providers do not always pay websites that allow them to collect data. ShareThis, for example, is a free service for websites to implement sharing functionalities. By putting 'Sharing Buttons' on their webpages, the sites technically allow ShareThis to inject their third-party cookies. See ShareThis, 'Privacy Notice' <http://www.sharethis.com/privacy> accessed 22 February 2017.

[60] For a detailed explanation of how data aggregators work, see Jeff Chester, 'Cookie Wars: How New Data Profiling and Targeting Techniques Threaten Citizens and Consumers in the "Big Data" Era' in Serge Gutwirth and others (eds), *European Data Protection: In Good Health?* (Springer 2012). For an exemplary list of data aggregators, see Ranker, 'The Top Advertising Data Exchanges and Aggregators' <http://www.ranker.com/list/advertising-data-exchanges-and-aggregators/online-ad-network-lists> accessed 22 February 2017.

[61] For technical details of how cookies are packaged and commoditised, see Dirk Bergemann and Alessandro Bonatti, 'Selling Cookies' (2015) 7(3) American Economic Journal: Microeconomics 259.

network.[62] Sometimes the information does not flow only in a one-way manner. These profiles might well feed on each other, creating one or more all-encompassing user profiles. It follows that a large part of one's online activities across the Internet are highly likely to end up contributing to numerous profiles — or one, complex, giant, all-encompassing, omnipotent 'database-in-the-sky'[63]. Some of them can be very revealing since a majority of the sites can be somehow connected to one profile by a common ad network provider, ad agent, data aggregator, or data provider. By breaking down the technical barriers of omnipresent tracking, the industry has virtually created an overarching cloud of profiling.

One more implication of cookie matching is that it facilitates tracking, because it allows one cookie setter to 'share' its recognition of a particular user to its business partners. That means, clearing the cookies of one website does not stop that website from recognising the same user in the future, since that website may learn from another website that this is a particular returning user that they once tracked. This leads to the creation of what are known as 'zombie cookies'[64] and makes tracking and profiling more powerful, and harder to avoid.

1.2.3 Targeting

Now with a complete profile of a user's interests, it seems something simple to decide what ads should be delivered to a certain user. However, this is in

---

[62] For a brief analysis of the value chain in the online advertising industry, see Sarunas Barauskas and Philippe Gondard, *Google : End Of The Online Advertising Bubble* (2016) <https://kalkis-research.com/google-end-of-the-online-advertising-bubble> accessed 27 October 2016.

[63] Paul Ohm, 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' (2010) 57 UCLA Law Review 1701, 1748.

[64] See Jonathan Mayer, 'Tracking the Trackers: Microsoft Advertising' (2011) <http://cyberlaw.stanford.edu/blog/2011/08/tracking-trackers-microsoft-advertising> accessed 13 March 2018.

fact much more complicated than it might sound. It is of course imperative to determine which kinds of advertisement a given user might show interest in, but there is much more involved than simply picking one of the advertisers in the most interesting category. The oversimplified understanding of the targeting phase of OBA comes partly from the terminology of 'targeting' itself.

The term 'targeting' might leave the impression that advertisements are categorised and then assigned to different groups of users that are locked into relevant interest categories based on their online profiles. This indeed used to be the business model at the very beginning of the history of OBA,[65] but it is no longer the case. Now the practice of targeting advertising should not be understood as a process of merely 'pushing' ads to viewers anymore, but one of 'pairing' viewers with advertisers on a real-time basis.

This is sometimes facilitated by a real-time bidding system.[66] The underlying idea of this system is that the opportunities to present advertisements to Internet users should not be sold as a package.[67] Instead, it is believed that each opportunity should be priced and sold separately, and instantaneously. This system works as follows: When a user is loading a webpage, an advertising slot is to be created on that page. The right to display an ad in this slot is called an 'impression'.[68] Impressions are open for sale in a way similar to an auction. To start the auction, the ad network provider broadcasts a message to all potential advertisers and invites them to participate in the bidding. Once the given time for bids runs out, the ad

---

[65] Martin Glanert, 'Emerging Trends in Online Advertising' (2010) <http://behavioraltargeting.biz/emerging-trends-in-online-advertising/> accessed 3 March 2018.
[66] Google, *Google White Paper: The Arrival of Real-Time Bidding, and What it Means for Media Buyers* (2011) <http://static.googleusercontent.com/media/www.google.com/en//doubleclick/pdfs/Google-White-Paper-The-Arrival-of-Real-Time-Bidding-July-2011.pdf> accessed 18 May 2016.
[67] Turow (n 57) 80-81; Smith (n 57) 60.
[68] Turow (n 57) 80-81; Smith (n 57) 15.

network provider selects a winner based on the price they offered.[69] It then notifies the successful bidder and requests the content to be shown in the advertising space. This process is completed on an entirely automated basis, and within a matter of milliseconds.[70]

One might probably wonder how the advertisers would decide whether an impression is worth bidding and how much they should offer. To ensure the bidding is rational, some information is sent around with the auction announcement. According to Google's technical document, such information may include the user's advertising ID, webpage URL, part of the IP address, browser version, postal code, geolocation, gender, age bracket, detected language, interest categories and their weights, device specifications and so on.[71] More importantly, cookie matching is also enabled in the process,[72] allowing advertisers to link the user to a profile in their own database. Based on all these details, advertisers can make an informed bidding decision.

Apparently, the real-time bidding process, as a means of targeting, amplifies the circulation of user data, and thus makes profiling even more ubiquitous. No statistical data is available on the number of bidders participating in the scheme, but considering the large part of online advertising Google is hosting,[73] the scale of information distributed every time a user loads a webpage would be conceivably enormous. Not all advertisers have the capability to take part in the bidding on their own, so many of them

---

[69] Turow (n 57) 80-81; Smith (n 57) 45-46.
[70] In the case of Google, the auction expires in 0.1 second. See Google, 'Get Started | Real-Time Bidding Protocol | Google Developers' <https://developers.google.com/ad-exchange/rtb/start> accessed 22 February 2017.
[71] Google, 'Real-Time Bidding Protocol Buffer v.138' <https://developers.google.com/ad-exchange/rtb/downloads/realtime-bidding-proto> accessed 13 March 2018.
[72] Google, 'Cookie Matching | Real-Time Bidding Protocol | Google Developers' <https://developers.google.com/ad-exchange/rtb/cookie-guide> accessed 22 February 2017.
[73] See section 1.3 below.

turn to bidding agencies.[74] It is true that with these agencies, the number of recipients of real-time data would be smaller. At the same time, however, it makes the profiling of these agencies more powerful as they bid on behalf of many websites and can therefore track more users across these sites. From that point of view, the conceptual relationship between tracking, profiling and targeting is not linear, but rather mutually supportive.

With an overview of online targeting in mind, it would be safe to conclude that the operation of the OBA industry resembles the business model of letting agencies (or dating websites) to some extent. A person works as an intermediary between the two sides of a potential transaction (or romantic relationship), usually holding the details of one side. Those who are interested from the other side may see the primary information of a potential match, but will have to pay for the chance to talk. All these similarities make the OBA industry more like an intermediary, except that the match is made at speed, and nearly always without the awareness of the ad viewers whose data is processed.

All three phases of the widely applied model of OBA, namely tracking, profiling and targeting, mirror the defining features of big data: high volume, variety and velocity. At the micro-level, an individual is now subject to the massive collection of behavioural and demographic data. Such data is generated from and circulated to a wide range of data points, sometimes even the offline ones.[75] The identification, representation and personalisation are all conducted on a quasi-simultaneous basis. At the macro-level, almost all Internet users are exposed to OBA, creating huge amounts of data being

[74] Google itself also runs a real-time bidding agency Invite Media. See Google, *Google White Paper* (n 66) 7.

[75] For instances of marketers collecting offline consumer data, see Daniel J. Solove, 'Privacy and Power: Computer Databases and Metaphors for Information Privacy' (2001) 53 Stanford Law Review 1393, 1408.

processed. Ad network providers, advertisers, publishers, ad exchanges, data brokers and bidding agencies are all contributing to the collection, transfer and utilisation of such data. Intense analytics and decision-making based on the data are achieved at a tremendous speed. As a result, everyone becomes part of a world-wide, enormous yet invisible network woven by all these actors. These statements will be further confirmed below from an economic angle.

1.3 Not just size matters: How powerful is the OBA industry?

1.3.1 Size — The dominant oligarchy

With a certainly over-simplified sense of how the routine communications of personal data are operating behind the screen, the discussion will now turn to the bigger picture. To begin with, a selection of financial figures are provided here. In the European advertising market, TV commercials had been the largest sector for a long time — until recently. According to a recent IAB Europe AdEx Benchmark Report, online advertising overtook the European TV market for the first time in 2015, with a market value of €36.4bn following an annual growth of 13%,[76] whereas in 1999, online advertising made up a mere 0.5% of the entire European marketing sector.[77] The robust growth of online advertising is also evident on a global scale. While TV advertising remained the largest category in the global advertising market at the time of writing, it is projected that digital advertising will take the lead in 2018 in

[76] IAB Europe and IhS, *AdEx Benchmark 2015: European Online Advertising Expenditure* (2016) 3 <http://www.iabeurope.eu/wp-content/uploads/2016/07/IAB-Europe_AdEx-Benchmark-2015-report_July-2016-V2.pdf> accessed 2 November 2016.
[77] IHS, *Paving the Way: How Online Advertising Enables the Digital Economy of the Future* (2015) <http://www.iabeurope.eu/files/9614/4844/3542/IAB_IHS_Euro_Ad_Macro_FINALpdf.pdf> accessed 28 October 2016.

terms of spend.[78] Depending on how impressions (the right to present ads to individual users) are bought and sold, the online advertising market can be further segmented into programmatic and non-programmatic (or mass buying). It is becoming the mainstream to conclude media transactions based on real-time bidding or similar technologies. A growing number of ad spots are now being traded with individualised, completely automatic mechanisms based on analyses of Internet user data. The vast majority of digital advertisers (92%), agencies (89%) and publishers (88%) in Europe have adopted programmatic advertising for display campaigns,[79] with more than half of their total spend now invested on programmatic advertising.[80]

Within a market of this size, a number of key players have been significantly dominant. A MAGNA report estimates that the annual revenues of the global digital advertising sector amounts to $209bn in 2017.[81] This includes a portion of (only) around 40% that goes to ad publishers (content providers), and the rest to the intermediaries like ad network providers and advertising agencies.[82] A large part of the overall revenues has been taken by

---

[78] Dentsu Aegis Network, *Global Ad Spend Forecasts* (2018) 7 <http://www.dentsuaegisnetwork.com/m/en-UK/DentsuAegisAdSpend/DANJune2018.pdf> accessed 8 June 2018.

[79] IAB Europe, *Attitudes towards Programmatic Advertising* (2016) 12 <http://www.iabeurope.eu/wp-content/uploads/2016/07/IAB-Europe-Attitudes-towards-Programmatic-Advertising-report_June-2016-v3.pdf> accessed 2 November 2016.

[80] IAB Europe and IHS, *European Programmatic Market Sizing 2016* (2017) 9 <https://www.iabeurope.eu/wp-content/uploads/2017/09/IAB-Europe_European-Programmatic-Market-Sizing-2016-report_FINAL-with-appendix.pdf> accessed 13 March 2018.

[81] MAGNA, *Global Advertising Forecast* (2017) 5 <https://www.magnaglobal.com/wp-content/uploads/2017/12/121117-MAGNA-Global-Forecast_Winter-Update_Final.pdf> accessed 8 June 2018.

[82] World Federation of Advertisers, *WFA Guide to Programmatic Media: What Every Advertiser Should Know about Media Markets* (2014) 7 <http://www.wfanet.org/media/programmatic.pdf> accessed 3 November 2016. See also Dempster and Lee (n 47) 35. This estimate also gains support from a 2017 ANA report, although a higher estimate is provided based on a different data set. See ANA and others, *Programmatic: Seeing Through the Financial Fog - An In-Market Analysis of Programmatic Media at the Transaction Level* (2017) 12 <http://www.ana.net/getfile/25070> accessed 13 March 2018.

the few oligarchs. Google's 2017 revenues from advertising services account for $95.38bn (which makes up 87% of the group's entire revenues).[83] Similarly, Facebook extracts most of its revenues from advertising, with $39.94bn (making up 98% of the total revenues) in the same year.[84] That means, Google and Facebook alone have grasped 65% of the whole online advertising industry's revenues.[85] Of course, it is open to dispute whether it is appropriate to compare self-reported figures with industrial estimates, but this would by and large give a sense of how a small number of behemoths are dominating the sector.

What makes the duopoly of Google and Facebook even more unsettling is the heterogeneous market structure and the unique value chain in the OBA sector. In the previous section, a business model has been outlined which involves a list of stakeholders: advertisers, ad publishers, ad network providers (ad exchanges), data providers, data aggregators and agencies. This is far from the level of complexity in the actual advertising market: First, the types of actors in the ad network are much more diverse than the depiction above; second, the consumption of marketing budget — as well as data — is not as linear as the example might make it seem. Including a much wider range of parties involved, Figure 1.1 below sketches out a (still very simplified) landscape of the OBA ecosystem. In the preceding section, some of these actors' roles have been briefly introduced with the made-up website names. Most of the categories shown in the diagram below can be roughly considered to be covered by the wider definitions of those actors: Exchanges are like ad

[83] Alphabet Inc., *Form 10-K* (2018) 28 <https://abc.xyz/investor/pdf/20171231_alphabet_10K.pdf> accessed 8 June 2018.
[84] Facebook Inc., *Form 10-K* (2018) 43 <http://d18rn0p25nwr6d.cloudfront.net/CIK-0001326801/c826def3-c1dc-47b9-99d9-76c89d6f8e6d.pdf> accessed 8 June 2018.
[85] It should be noted that Google and Facebook are acting both as ad publishers and intermediaries in the online advertising market.

networks (*ads.com*); retargeting services and data suppliers are data providers (*data.com*); DMPs (data management platforms), tag management systems and AMPs (audience management platforms) are like data aggregators (*aggregate.com*); and agency trading desks and DSPs/SSPs (demand/supply-side platforms) are like advertising agents (*agent.com*). It is not intended to explain the roles of all these participants here, and many of them are not necessarily engaged in every single piece of advert displayed to a user. Suffice it to know at this point that most of them are part of the sophisticated decision-making process that brings an advertiser to an ad viewer. It is also noteworthy that in the OBA market, data selling is not just about 'handing over' the user data to the buyer; it is more about allowing the buyer to set and read cookies directly on the user device (by means of cookie matching). That is exactly why when a user visits *Esquire*'s homepage, a total of 18 advertising-related services would be informed:[86] The right to place cookies on the terminal device is sold to a different service, who keeps a record and resells the right to one of its interested customers, who again sells the right to someone else. Layers of transactions of the right all take place in a few thousandths of a second.

[86] Smith (n 57) 23.

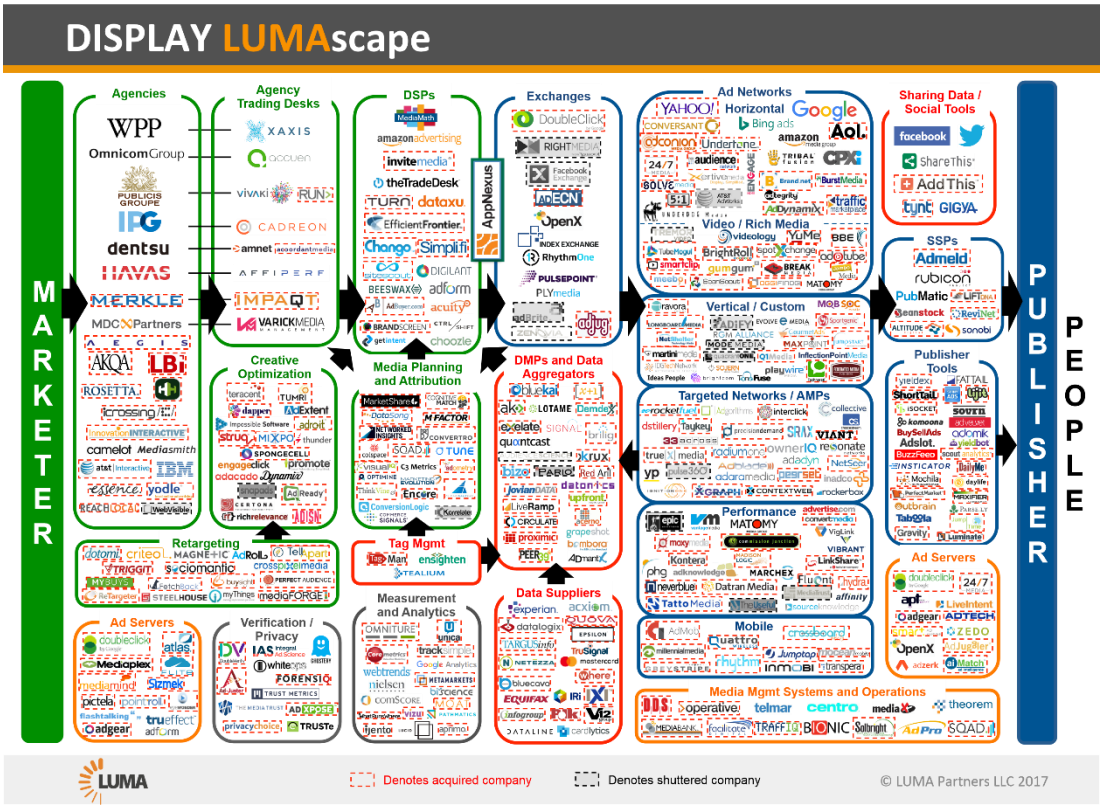

Figure 1.1 Display LUMAscape[87]

The heterogeneity of the players in the market, combined with the business model of selling the 'right to place cookies', means that the competitors in this market are not only competing against, but also enabling, each other. A new player's entry into the market may probably dilute the market share of an existing player, but the latter may end up gaining access to more Internet users through the complex network that 'sells' cookies from one player to another. It follows that, in a big data world, as long as a business has sufficient technological capacity and access to the data network, it can reach an audience much bigger than what its market share might suggest.

The power of a company to seize online behavioural data grows disproportionately as its size does. A research project spanning 18 months has

[87] LUMA, 'Display LUMAscape' <https://www.lumapartners.com/luma-institute/lumascapes/display-ad-tech-lumascape/> accessed 13 March 2018. The author would like to thank LUMA Partners LLC for the permission to use the graphic for this study.

just provided the latest empirical evidence to support this hypothesis. The researchers developed a web crawler to collect information about 3.2 million cookies employed by the most popular 100K websites. It is found that, among all the services that have access to third-party cookies, the most powerful ones (<1%) have actually covered 75% of the websites across the Web.[88] Another finding is that, for all the third-party cookies (which significantly outnumber first-party cookies), doubleclick.net — a domain name owned by Google for its OBA business — tops the chart by having cookies on 42.1% of all websites.[89] In other words, Google has access to information about an average Internet users' activities on more than 40% of the websites that they might visit. Another latest study focused on the mobile app sector provides an even more striking figure: 88.44% of the Play Store apps contain trackers from Google subsidiaries.[90] While probably not each of these websites or apps has a direct partnership with Google, what it takes for the latter to benefit from the former's connection with its users is simply a remote relationship within the ad network where one of the publisher's partners — be it agencies, data providers, platform providers or whatsoever — or *their partners* takes part in the network.

1.3.2 Breadth — The multiple arms of the businesses

Sitting in the centre of the advertising network that brings data, money and sub-networks, Google is already controlling a massive amount of resources, but its expansion in the sector has yielded even greater powers. In the online advertising industry, Google plays a role that is much more than a mere

[88] Aaron Cahn and others, 'An Empirical Study of Web Cookies' (Proceedings of the 25th International Conference on World Wide Web, Montréal, 11-15 April 2016) 891.
[89] ibid 896.
[90] Reuben Binns and others, 'Third Party Tracking in the Mobile Ecosystem' (Proceedings of the 10th ACM Conference on Web Science, Amsterdam, 27-30 May 2018).

platform provider. Indeed, Google has never stopped expanding its business to further corners of the ecosystem. The starting point of looking into how far-reaching Google actually is rests in the fact that its business model is built on analysing user metrics and behaviour and selling such insights to advertisers.[91] It is also helpful to recall that the ultimate fuel fed into the enormous advertising engine is data, which enables marketing firms to optimise the delivery of adverts. With these two points in mind, some seemingly non-advertising-related products of Google would reveal their potentials in creating lucrative opportunities for the advertising giant.

Think of Google Maps, which stores the locations a user searched for, reviewed, saved and shared.[92] It is not difficult to imagine how profitable it would be for Google to have even just a not-so-accurate sense of the locations a user is looking for. An even more precise tracker would be Google Maps on a smartphone, which tracks not just where a user is heading, but also where exactly the user is. Now the mobile phone operating system, Android, might have already crossed your mind. Again, the overwhelming dominance in the smartphone market certainly opens the door to the extraction of further values from as much user data as they can collect. These include Gmail, YouTube, Chrome, Google Home, Self-Driving Car and so on. Whether those products that have already won global popularity or those that are still under development, almost every time Google reaches a new niche, it opens up a further ground for it to invest in, and then harvest, data for its marketing business.

---

[91] Stuart Sumner, *You: For Sale: Protecting Your Personal Data and Privacy Online* (Elsevier 2016) 77.

[92] Google, 'Delete Directions and Places from Your History' <https://support.google.com/maps/answer/3137804> accessed 13 March 2018.

Apart from developing new products or services, Google is also ready to exercise its capital power: buying along the value chain. Here, it is worth revisiting how Google has grown to its current size with a very brief timeline. As is well-known, Google started as an Internet search engine service in 1996.[93] Four years later, in 2000, Google started to move into the online advertising industry by launching AdWords, a service that allowed advertisers to purchase text-only ad slots on Google's search result pages based on relevant search terms.[94] At that point, Google acted only as an ad publisher. In 2002, Google introduced an auction feature into AdWords, by which advertisers may 'bid' on certain keywords and then winner's ads would be shown along search results for those specific terms.[95] This way, Google had in effect turned itself into also a platform, a demand-side platform (DSP), although the only inventory (ad space) supplier remained itself. In June 2003, Google rolled out a new service named AdSense,[96] shortly after the acquisition of Applied Semantics, a company with technologies to 'enable[] web publishers to understand the key themes on web pages to deliver highly relevant and targeted advertisements.'[97] By then, Google had become both a DSP (with AdWords) and an SSP (supply-side platform, with AdSense). In the same year and the following, Google purchased a number of services that seemed unrelated to its advertising business, including Blogger, Picasa and Google Earth.[98] In 2005, Google acquired Urchin, whose technologies formed the

[93] Google, 'Our History in Depth' <https://www.google.co.uk/about/company/history> accessed 21 December 2016.
[94] ibid.
[95] John Battelle, *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture* (Nicholas Brealey Publishing 2005) 142.
[96] ibid 151-152.
[97] Google, 'Google Acquires Applied Semantics' (2003) <http://googlepress.blogspot.co.uk/2004/04/google-acquires-applied-semantics.html> accessed 13 March 2018.
[98] Google, 'Our History in Depth' (n 93).

foundation of Google Analytics, perhaps one of the most influential products of Google.[99] It was later improved after the acquisition of blog analytics tool Measure Map.[100]

Google Analytics turned out to be such a popular tool for website operators to monitor their traffic that over 64% of the top 20 million most popular websites have now deployed the service.[101] This is particularly useful for ad publishers as it offers certain features that are specially designed for websites running ads.[102] However, for those websites that do not include ads or opt out from such advertising features, Google Analytics would nevertheless collect user data and might use such data for advertising purposes. [103] The introduction of Google Analytics has therefore significantly empowered Google in the online marketing sector as it gains greater capacity to measure advertising performance as well as to obtain online user data. In 2006, Google made another ambitious move of acquiring YouTube.[104] Since then, Google has become not only a prevailing publisher of video content,[105] but also a powerful video ad publisher, as well as a giant collector of entertainment behavioural data. To further expand its advertising kingdom, Google acquired DoubleClick in 2008. DoubleClick was well-known for its outstanding expertise in real-time ad-serving for publishers,[106] which was an indispensable

[99] ibid.
[100] Google, 'Here comes Measure Map' (2006) <https://googleblog.blogspot.co.uk/2006/02/here-comes-measure-map.html> accessed 22 February 2017.
[101] Datanyze, 'Web Analytics Market Share Report' <https://www.datanyze.com/market-share/web-analytics> accessed 13 March 2018.
[102] Google, 'About Advertising Features' <https://support.google.com/analytics/answer/3450482> accessed 13 March 2018.
[103] Google, 'How Google Uses Data When You Use Our Partners' Sites or Apps' (n 27).
[104] Google, 'Our History in Depth' (n 93).
[105] Randall Stross, *Planet Google: One Company's Audacious Plan to Organize Everything We Know* (Free Press 2008) 109-128.
[106] Smith (n 57) 79-83.

asset for Google to start running its own Ad Exchange in 2009.[107] This was helped by its acquisition of real-time bidding agency Invite Media in 2010,[108] enabling Google to further improve the performance of the real-time bidding system and further share the market revenues. In the same year, AdMob, a mobile advertising company, was acquired as part of Google's broader mobile strategy.[109] All of these takeovers show how Google has attempted to maximise its influence over many sectors, within or outside the advertising industry, by buying out potential competitors. This has remarkably bolstered Google's power not just economically, but also in terms of data use: In 2012, Google decided to merge user data across services — including YouTube, Gmail, Google Maps, Google+ and Android — so user activities on these services can be better integrated for targeted advertising purposes.[110]

To be fair, Google is not the only one who buys along the value chain. Other powerful players in the area have also made important moves towards a new, wider landscape with, of course, greater amounts of data: Facebook towards Atlas's ad server technology, LinkedIn towards Bizo's business audience marketing technology, AOL towards marketing optimization platform Convertro and Twitter towards mobile exchange MoPub.[111] Now it is almost impossible to find a pure platform provider in the advertising ecosystem. In order to the improve the performance of a network, data is indispensable, and

---

[107] Google, 'The DoubleClick Ad Exchange: Growing the Display Advertising Pie for Everyone' (2009) <https://googleblog.blogspot.co.uk/2009/09/doubleclick-ad-exchange-growing-display.html> accessed 22 February 2017.
[108] Seth Weintraub, 'Google Continues Buying Spree, Picks Up Invite Media' *Fortune* (2 June 2010) <http://fortune.com/2010/06/02/google-continues-buying-spree-picks-up-invite-media/> accessed 13 March 2018.
[109] Google, 'We've Officially Acquired AdMob!' (2010) <https://googleblog.blogspot.co.uk/2010/05/weve-officially-acquired-admob.html> accessed 27 May 2010.
[110] The Guardian, 'Google User Data to Be Merged Across All Sites Under Contentious Plan' *The Guardian* (25 January 2012) <https://www.theguardian.com/technology/2012/jan/25/google-merge-user-data-privacy> accessed 18 March 2018.
[111] Dempster and Lee (n 47) 18.

the easiest way to obtain valuable and renewable data is obviously to buy an undertaking that has a trove of valuable data or the experience in collecting and refining data.

1.3.3 Impact — The effectiveness of online marketing

The increasing market share of programmatic trading within the online media sector as well as the relentless investment on new technologies facilitating its operation is the best proof of the commercial success of OBA. One might naturally wonder about the extent to which OBA can help marketers turn their costs on the state-of-the-art infrastructures and layers of intermediaries into measurable profits. Or is it just yet another dot-com bubble? Although there is empirical evidence suggesting that OBA can boost productivity, the exact degree of effectiveness remains inconclusive. In 2008, a research team led by Microsoft Research Asia conducted an experiment-based study, which showed that behavioural targeting can improve the click-through rate (CTR) of all the covered ads by as much as 670%, if the optimal algorithm being tested is applied.[112] The methodology in that study was criticised by researchers from Yahoo! and Stanford University, who pointed out the potential existence of selection bias in the research, that is, the suspicion that the group of targeted users has a higher CTR not because of the relevance of the advertised product, but that those users are more likely to click through *any* generic ads.[113] However, even with such effects accounted for in their calculation, the team still saw a rise of 79% in CTR.[114] Advanced technologies adopted by

[112] Jun Yan and others, 'How Much Can Behavioral Targeting Help Online Advertising?' (Proceedings of the 18th International Conference on World Wide Web, Madrid, 20-24 April 2009).
[113] Ayman Farahat and Michael Bailey, 'How Effective is Targeted Advertising?' (Proceedings of the 21st International Conference on World Wide Web, Lyon, 16-20 April 2012).
[114] ibid 119.

advertisers, publishers, network providers and other supporting entities, such as real-time bidding, have made the improvement even more significant. According to Google's 'internal data', in a comparison of campaign performance between traditional and real-time bidding mechanisms on Google Ad Exchange in 2011, real-time bidding managed to save 19% on costs and raise CTRs from 0.09% to 0.15%.[115] It should be noted that these figures come largely from research projects with affiliation to major OBA network providers like Google and Microsoft, who are probably the only ones who have the first-hand data. That said, from a practical perspective, the popularity of OBA in the online marketing sector may serve as an indicator of its better performance, albeit possibly more or less exaggerated. According to Turn, a marketing platform for advertisers, advertisers who have adopted real-time bidding are seeing up to 135% improvement on CTRs and 150% on conversion rates (percentage of users who eventually make the purchase). [116] An experienced marketer claims that efficient online marketing informed by data-driven insights may boost addressable impressions by more than 400% and cut average costs per conversion by 30 to 40% for the better performance.[117]

The efficiency of OBA results largely from its ability to predict the 'right' group of ad viewers that are more likely to click through and make a purchase. Such predictions rely much on the intense tracking, profiling and targeting practices based on behavioural data, and it is not hard to imagine that the bigger the data pool is, the more useful the insights would be. This is exactly where big data connects the technological world with the commercial one: In the context of OBA, big data means big business, and big insights mean big

---

[115] Smith (n 57) 64-65.
[116] PubMatic, *Understanding Real-Time Bidding (RTB) From the Publisher Perspective* (2010) 11 <http://pubmaticblog.com/wp-content/uploads/2014/06/Understanding_RTB_Q12010.pdf> accessed 20 March 2015.
[117] Dempster and Lee (n 47) 10.

profits. As one commentator puts, 'Google is better because it's bigger, and it's bigger because it's better.'[118] The online advertising industry is a scale business that favours big entities,[119] because a firm's ability to process data grows exponentially as its size does, and the effectiveness of online marketing grows exponentially as data size does. The more Google knows about individual Internet users, the more effective their advertising services can be.[120] The internal drive characterised by the cycle of more efficient targeting, greater conversion and more effective remarketing leads to a richer customer portfolio.[121]

For this reason, the dominant position of Google in the market can hardly be challenged. When faced with the accusations of monopoly in certain markets, Google's standard response has always been that barriers to entry into the Internet market are low, and that any start-ups may copy the success of Google, just like how it took down Yahoo! and other competitive players.[122] This argument might be valid in around 2000, but does not seem so today anymore, not at least in the area of online advertising. Possession of the most critical resources in the industry, access to individual demographic and behavioural data, will only be more concentrated to those who are already rich in those assets. New services targeting a particular niche market might nonetheless be able to emerge, but to make it a big success, it has to rely on, and eventually be part of, a larger network that has been woven by the big players. To such an extent, while size is not the only decisive factor when it

[118] Siva Vaidhyanathan, *The Googlization of Everything (and Why We Should Worry)* (University of California Press 2011) 20.
[119] IHS (n 77) 13.
[120] Vaidhyanathan (n 118) 83.
[121] Dempster and Lee (n 47) 10.
[122] Vaidhyanathan (n 118) 19.

comes to the evaluation of market power in the OBA sector, it does matter, and it does help certain dominant businesses grow faster at other dimensions.

Now that we have seen the market power in terms of the size, breadth and impact of the dominant players in the OBA ecosystem, what does all that mean to individual Internet users? To bring the pieces together and make sense of the bigger picture throughout this section, we need to recall the landscape that has been sketched out, with the help of Figure 1.2.

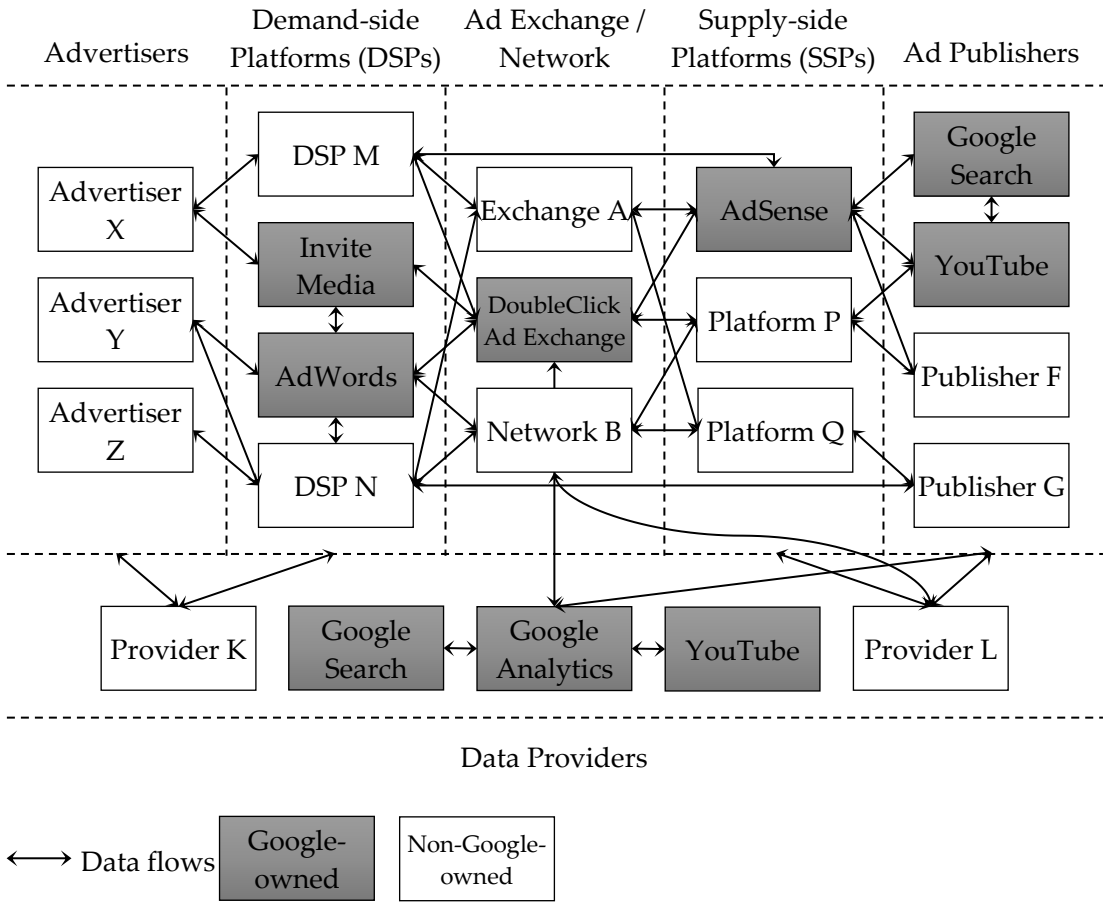

Figure 1.2 An OBA ecosystem

This (very simplified) diagram shows that Google actually operates services across most of the links within the OBA value chain. Each player in this ecosystem has connected to one or more upstream or downstream players, and they exchange user behavioural data (by giving data directly, informing decision-making or allowing one another to set cookies on user device) as part

of their partnership. There are of course other players who are not connected or only remotely connected to this network, but they would only take up a small portion of the entire industry. More importantly, despite the claim that entry barriers in this sector are relatively low,[123] for a start-up to survive and grow in this data-driven industry, they will have to be part of this network to seek either suppliers or customers, or probably both, or else they will become marginalised. This would have some unsettling implications from a competition or consumer protection point of view. If the entire ecosystem has eventually become *one* network, the mere increase of competitors will only strengthen the power of the already powerful, at least in terms of data power. New entrants will bring in new ways to collect, refine, share or analyse online data, which will end up feeding into the system and benefit the existing players. As such, market share would not be a good indicator of power in this realm. Also, the myth that more choices in the market can promote consumer welfare[124] would apply at best only partially here. For example, an online video viewer in the set-up of Figure 1.2 might decide to switch from YouTube to a non-Google service G, who subscribes to Q as its SSP. Platform Q sells G's inventory (advertising spaces) to network providers A and B, neither run by Google but B happens to be one of AdWords' networks. By means of cookie matching, Google will still be able to place a cookie on the viewer's computer, even though they have moved from YouTube to the only-remotely-connected-to-Google service G. Granted that online users do have the choices of a range

[123] Miguel Helft, 'Google Makes a Case That It Isn't So Big' *The New York Times* (28 June 2009) <http://www.nytimes.com/2009/06/29/technology/companies/29google.html> accessed 13 March 2018.

[124] Commission, 'Communication from the Commission — Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings' (2009) OJ C 45/7, para 19.

of services, they would remain with no choice in respect of the giant advertising network behind those services.

1.4 Summary: That is not the Google I know (anymore)

Back in the 1960s, data storage was so limited and costly that computers were designed to use only two digits to indicate the year — which led to what was known as the 'Millennium Bug' later in the 1990s.[125] Today, this is not the case anymore. Data processing, including storage, transfer and analysis, has become so enormously affordable that the costs of keeping as much data as possible are negligible, compared to the potential value that might derive from the data. As economic constraints cease to stand in the way of the effective use of data, massive digitalisation of almost everything has become the norm.[126] The ability to handle large amounts of data from a broad scope of information feeds at tremendous speed has changed not just technological configurations, but also the societal ones. The application of big data to OBA shows perfectly how the data super-powers have made the most of this trend to maximise their influences, and profits. From jostling for sectoral dominance to reaching outwards across the ecosystem, and then securing market impact, key players in the industry have never spared any effort to make themselves more ubiquitous. Technical solutions to such an end have also evolved beyond the understanding of average Internet users and sometimes even regulators. They might have an idea of the existence of OBA practices, but most of them might probably feel surprised had they learned about the behind-the-scenes details of how these systems actually work.

---

[125] Feng Li, Howard Williams and Martin Bogle, 'The "Millennium Bug": Its Origin, Potential Impact and Possible Solutions' (1999) 19(1) International Journal of Information Management 3.

[126] Viktor Mayer-Schönberger, *Delete: The Virtue of Forgetting in the Digital Age* (Princeton University Press 2009) 52.

This chapter is intended to offer a closer but different perspective to understand the reality of the OBA industry. An OBA network typically has a tremendously complex architecture, in both technical and organisational terms. As massive data travels from individuals to different types of intermediaries, and between a multitude of them, the marketing values emerge, and become carved up among these players. The system is indeed smart, yet in a sense different from the way many would see it. It is not necessarily designed to read a user's mind; rather, its aim is probably only to maximise the output of commercial campaigns with minimised costs, and then split the revenues among those actors. That statement holds true for Google: It is probably neither morally good nor evil. At the end of the day, it is just a profit-seeking advertising company.[127] Along with the routine, invisible cycle of tracking, profiling and targeting that every single Internet user would experience hundreds or even thousands of times a day, big business interests keep growing in size, breadth and impact. Yet, as will be shown in the subsequent chapters, these mundane activities have in fact caused a great deal of economic, social, legal and political consequences.

To have a fuller understanding of the big data phenomenon, the next two chapters will further investigate the broader implications of big data in contexts beyond the techno-economic landscape. The findings in this chapter will provide a useful basis for making sense of both the potential benefits and the risks of OBA as an instance of big data. The arguments advanced in the ongoing debates surrounding various dimensions of OBA — whether for or against its intensive use of data — will be examined in a critical manner. The legitimate interests as well as the positive values mainly claimed by the

[127] In its latest financial report, Google (or more precisely, Alphabet) identifies itself as an incorporation that 'generate[s] revenues by delivering online advertising that consumers find relevant and that advertisers find cost-effective'. See Alphabet Inc. (n 83) 3.

marketing industry will be discussed in Chapter 2, whereas the possible harms — whether individualistic, collective or societal, and whether immediate or intangible — arising from the use of data for OBA purposes will be theorised in Chapter 3.

Chapter 2 Big Promises That Big Data Holds: Legitimate Interests and Societal Benefits in the Context of Online Behavioural Advertising

The triumph of big data in many sectors, as least in business terms, does not come with no reason. In fact, big data is considered by many as a revolutionary phenomenon that holds a wide range of promises. There should be little disputing that big data has changed and will continue to change the life in a highly digitalised world. Much less visible, however, is the potentially world-shattering reconfiguration of the powers and interests involved in the shifted paradigm of materialising the values of data. It is therefore worth a thorough inquiry into how a variety of stakeholders will be impacted by the rise of big data, and whether such impacts are justifiable.

Based on the developments examined in the previous chapter, a number of major claims over the legitimacy of OBA practices can be looked at more closely now. The arguments can be examined loosely under three strands of discussions, respectively from an individualistic, economic and societal perspective.

2.1 Good for individuals? Relevant ads and free content

2.1.1 Personalised advertising and user preferences

(a) Contradictory empirical evidence

A point constantly made by the marketing industry is that OBA brings Internet users more personalised, tailored and interesting advertisement, and hence improves their online experiences. To prove this point, they have sponsored a number of surveys asking users about their perceptions of online advertising

practices. In 2013, for instance, the Digital Advertising Alliance (DAA) commissioned Zogby Analytics to carry out a survey with a sample of 1,000 US respondents.[1] In this survey, a participant was first asked a number of questions regarding the importance of free online content, their preferences of free or paid content, and the utility of online ads.[2] The sixth question reads as follows: 'Would you rather see Internet ads for random/generic products and services, or ads for products and services that reflect your interests?'[3] 16.1% of the surveyed said they would prefer the former option, 40.5% the latter and 27.6% both.[4] Based on this seemingly significant margin, a trade organisation claims that American Internet users are 'largely comfortable with the value-for-value exchange that interest-based advertising represents.'[5] A different piece of supporting evidence came from a 2009 survey sponsored by TRUSTe and independently conducted by TNS.[6] The study found that 71.7% of the 1,008 respondents strongly or somewhat agree that 'online advertising [is] intrusive and annoying when the products and services being advertised are not relevant to [their] wants and needs.'[7] A similar study was carried out by Westin in the same year, which found out that, assuming four stringent

[1] Zogby Analytics, *Interactive Survey of US Adults* (2013) <http://www.aboutads.info/resource/image/Poll/Zogby_DAA_Poll.pdf> accessed 16 March 2017.
[2] ibid 1-5.
[3] ibid 6.
[4] ibid.
[5] Network Advertising Initiative, 'US Internet Users Understand and Value Interest-Based Advertising, According to DAA Survey' (2013) <http://www.networkadvertising.org/blog/us-internet-users-understand-and-value-interest-based-advertising-according-daa-survey> accessed 19 March 2017.
[6] TNS, *2009 Study: Consumer Attitudes About Behavioral Targeting* (2009) <https://dsimg.ubm-us.net/envelope/104162/339732/1249993418527_TRUSTe_TNS_2009_BT_Study_Summary.pdf> accessed 20 March 2017.
[7] ibid 5.

privacy policies are adopted, most respondents (except for the 63+ age group) would feel 'comfortable' with behavioural targeting and tailoring.[8]

These statistical figures quoted in industry-led studies are however subject to quite heavy criticisms. In 2009, for example, Turow et al conducted a series of telephone interviews with a nationally representative sample of 1,000 US adults.[9] This study was carried out to verify previous research findings that are allegedly in favour of the industry, including the TRUSTe-TNS one, as well as the one led by Westin.[10] The report questions the methodology employed in both surveys for a number of reasons.[11] First, both surveys are run in the form of online questionnaires and on a voluntary basis, which might lead to the under- or even non-representation of those who have had privacy concerns and thus have not participated; second, the questions in the surveys are believed to be too general, having intermingled the two issues of ads being tailored and ads being tailored *based on a particular tracking device*; third, the nature of information used in online advertising is not clearly defined in the questionnaire. [12] Against these drawbacks, the Turow-led study has incorporated a number of improvements, including the use of random digit dial to reach interviewees and more specified questions.[13] They found that 'fully 66% of the respondents do not want advertisements tailored for them'.[14] This was later confirmed by a separate telephone survey conducted by Pew

---

[8] Cited from Joseph Turow and others, 'Contrary to What Marketers Say, Americans Reject Tailored Advertising and Three Activities That Enable It' (2009) <http://ssrn.com/abstract=1478214> accessed 20 February 2015. This study has not been published in any scholarly journal and is now only available in secondary sources. See Chris Jay Hoofnagle and Jennifer M. Urban, 'Alan Westin's Privacy *Homo Economicus*' (2014) 49 Wake Forest Law Review 261, 264-265.

[9] Turow and others (n 8).

[10] ibid 9-11.

[11] ibid 10-11.

[12] ibid.

[13] ibid.

[14] ibid 14.

Research Center in 2010 on a larger sample of 1,729 adult Internet users: 68% responded 'I'm NOT okay with it because I don't like having my online behavior tracked and analyzed'.[15]

Yet, the findings above are also faced with counter-criticisms. A study by Sableman et al in 2013 challenged the transferability of the findings by Turow et al to the context of policymaking.[16] The weaknesses of the study, as Sableman et al argued, rest in the inclusion of non-users of the Internet, and the reliance on merely hypothetical questions.[17] It was suggested that a meaningful investigation should be based on real-life scenarios. Hence, in this survey on 150 students from a university, Sableman and co-authors designed a two-part process. They first asked respondents how they felt about five scenarios where the setups were different in terms of being online/offline, personalised/non-personalised and news/social media/retailer sites.[18] For example, a participant might be asked to imagine they are an avid photographer, and then to rate how they would feel about seeing ads of camera-related products on CNN.com.[19] In the second part, the respondents were asked the same questions regarding their 'privacy' attitudes as in previous studies, including Turow et al's.[20] They found in the first part that 53% of the surveyed users felt positive about online tailored ads. When it comes to the result of the second part, the numbers fell back to a level comparable to

[15] Kristen Purcell, Joanna Brenner and Lee Rainie, *Search Engine Use 2012* (2012) 23 <http://www.pewinternet.org/files/old-media//Files/Reports/2012/PIP_Search_Engine_Use_2012.pdf> accessed 1 December 2016.
[16] Mark Sableman, Heather Shoenberger and Esther Thorson, 'Consumer Attitudes Toward Relevant Online Behavioral Advertising: Crucial Evidence in the Data Privacy Debates' (2013) 2013(1) Media Law Resource Center Bulletin 93.
[17] ibid 101-102.
[18] ibid 104-105.
[19] ibid.
[20] ibid.

previous studies.[21] It was therefore suggested that Internet users might respond differently to scenario-statements and belief-statements, and that consumers' 'needs and preferences must be correctly understood and balanced.'[22] This study, however, has clearly not overcome the common pitfall in previous research projects that have mixed up the idea of 'personalisation' with that of 'tracking'. In fact, they have even included an offline scenario in the survey as 'a foil to prevent the respondents from guessing the manipulation'.[23] The author seems to suggest that, when given too much information about the privacy implications of OBA, respondents are more likely to become irrationally over-sensitive.

(b) Defects of existing findings

These seemingly contradictory conclusions drawn from different studies have proved a rather obvious fact that the self-reported attitudes of Internet users towards OBA practices depend heavily on the exact wording of the questions. When such terms as 'tailored', 'personalised', 'relevant' or 'interests' are used to represent OBA, the users tend to feel more positively about it, whereas the attitudes tend to be much more negative when questions are rephrased with terms like 'tracking', 'targeted', 'monitor', 'privacy' and so on. This however should not come as much of a surprise as the phenomenon of 'privacy paradox' has been well-documented in the literature: When it comes to privacy concerns, user behaviour and statements are often dramatically inconsistent.[24] Here, two

[21] ibid 105-107.
[22] ibid 107-108.
[23] ibid 107.
[24] See, for example, Naveen Farag Awad and M. S. Krishnan, 'The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to Be Profiled Online for Personalization' (2006) 30(1) MIS Quarterly 13; Patricia A Norberg, Daniel R Horne and David A Horne, 'The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors' (2007) 41(1) The Journal of Consumer Affairs 100; H. Brian Holland, 'Privacy Paradox 2.0' (2010) 19 Widener Law Journal 893.

problems will severely impact the credibility of empirical studies in this area, one concerning methodology and the other relevance.

The methodological problem concerns whether the questionnaires are designed in a professional, objective and neutral way, which should be tested against the standards of social statistics. It is of paramount importance that researchers have asked the right question. Factors that might impact the impartiality of their findings may include, for instance, what are the typical knowledge level with which the users' preferences should be gauged? Should it be an 'average user' with only a reasonable level of knowledge about OBA? Or should it be an 'informed user' who is familiar with the technical details as well as the arguments from both sides? Also, what information should be given to a respondent before they take the survey? Some researchers believe that respondents are supposed to be better-informed of the technical details prior to the questions;[25] others, like Sableman et al, on the contrary, maintain that the real-life reactions of Internet users are what matters, and excessive information would only be manipulative. From a methodological point of view, a robust, policy-oriented investigation should provide impartial information about both the risks and the benefits of OBA beforehand. This is because the role of a policymaker is to facilitate individuals to act in accordance with their informed choice, and also if such a sensible choice proves impractical on an individual level, to make such a choice on behalf of them as a collective. Unfortunately, it seems that none of the studies presented above has sufficiently highlighted the main arguments put forward by *both* sides.

[25] See, for example, Edith G. Smit, Guda Van Noort and Hilde A.M. Voorveld, 'Understanding Online Behavioural Advertising: User Knowledge, Privacy Concerns and Online Coping Behaviour in Europe' (2014) 32 Computers in Human Behavior 15.

The problem of relevance — perhaps the even more fundamental one — concerns the extent to which it is appropriate to apply such empirical conclusion to inform policymaking. The real question here could be the relevance of Internet users' attitudes. In other words, how much should the findings about the general perceptions of OBA within a territory be taken into account when making decisions about a regulatory framework at all? All this comes down to one underlying question: Is it justifiable to adopt a policy on how the OBA industry may make use of personal data based on the fact that the majority of the society prefer tailored ads or not? If online marketers truly believe that the majority of the population would prefer ads relevant to them, they should leave the decision to individuals. As long as it is technically and economically feasible, the choice of those who do not wish to be served with personalised ads should be respected, even if they make up only a marginal part of the population. If it is true that most Internet users would prefer tailored ads, then online marketers should not be worried about giving users the choice. If they are, it is more likely that most users simply do not actually prefer tailored advertising, or at least do not care. The fact that many users do not opt out targeted advertising may simply be a consequence of inertia or ignorance, not that they actually like it. Either way, the mere claim that a high percentage of Internet users indeed want tailored advertising would not sufficiently justify the assumption that *everyone* does.

2.1.2 Free content and services financed by advertising revenues

Apart from presenting users with relevant ads, trade groups on behalf the OBA sector often also make another argument that individual consumers benefit much from today's online advertising, because a large part of online services and content are financially supported by an advertising-based business model. To make this claim sound better-founded, again, a

considerable number of studies have been launched in their quest for substantial evidence. Such initiatives have largely focused on two supporting statements that, subjectively, Internet users would rather trade their information than pay in monetary forms for online content, and that, objectively, they become better off in economic terms.

(a) Consumer preferences

The first prong of the quest would inevitably involve surveys again, as the aim is to investigate the personal preference of individual users. Not surprisingly, when confronted with the two options of being tracked, profiled and targeted, and being charged, a significantly larger portion of people are found ready to go for the first option without much hesitation. The 2013 Zogby Analytics survey referred to in the previous section, for example, also included questions asking respondents to pick a side when they could not have it all. One question reads as follows: 'Which of the following would you prefer: an Internet where there are no ads, but you would pay for most content like blogs, entertainment sites, video contact and social media, or today's Internet model in which there are ads, but most content is free?'[26] 75% of the surveyed opted for the latter, and only 9.3% for the former.[27] A similar survey was conducted in 2016, also funded by the DAA, including the same question and same result by an even bigger landslide of 85.2% in favour of ad-supported free content.[28] The relative difference in the outcomes might have resulted from the way the 2016 questionnaire was designed. Before answering the questions, the participants were asked to evaluate how much they think one would have to pay each

[26] Zogby Analytics (n 1) 2.
[27] ibid.
[28] Zogby Analytics, *Public Opinion Survey on Value of the Ad-supported Internet* (2016) 5 <http://digital.daaoperations.org/sites/digital.daaoperations.org/files/DAA_files/ZogbyAnalyticsConsumerValueStudy2016.pdf> accessed 16 March 2017.

month respectively for a total of 17 categories of online services.[29] These services range from search engines (with an average estimate of $4.46) to video streaming providers ($7.83).[30] With the potential costs added up in mind, the respondents may presumably have become more sensitive to pecuniary incentives.

However, as with the surveys conducted to quantify user perceptions of tailored ads, these studies of consumer benefits are also methodologically flawed in several aspects. First of all, the information given to the surveyed is noticeably incomplete, particularly skewed towards the potential benefits of online advertising. It is debatable whether asking the respondents to evaluate the prices of such services should be regarded as 'manipulative' or 'informative'. In any event, asking only about the price but not the cost is certainly partial, as the focus would be put unevenly on only one end of the equation. As Hoofnagle and Whittington have illustrated, the disproportionate emphasis on the price rather than the cost in an online setting may put consumers in a vulnerable position.[31] Again, the lack of empirical studies that aim to present impartial information for both sides has left a regrettable gap in the literature. The second flaw of these studies is the unclear distinction between different forms of online advertising. For Internet users, the 'cost' of non-tailored advertising is in theory lower than that of OBA based on multiple tracking, profiling and targeting devices. Therefore, it is conceivable that some respondent might have reacted differently had it been made clear what kind of online advertising is concerned and what techniques are involved. Third, the options offered in the questionnaires are very limited

[29] ibid 1-4.
[30] ibid.
[31] Chris Jay Hoofnagle and Jan Whittington, 'Free: Accounting for the Costs of the Internet's Most Popular Price' (2014) 61 UCLA Law Review 606.

in that only two scenarios have been envisaged: a world where advertising replaces the charges for most online services, and a world where all these services are fully funded by user subscription. What the surveys do not make the respondents realise is that there might be a world where individual users may choose to accept ads or to opt for an ad-free version for a fee on a service-by-service basis. Given the technical feasibility of such a third option, the deprivation of such a possibility can only be considered unreasonable and may have caused many of those who would have opted for the third approach to end up opting for the one with ads. Besides all these methodological defects, of course the relevance criticism outlined in the previous section might well be applicable to this context, too. The fact that the majority of Internet users do prefer paying for services in the form of personal data cannot justify the deprivation of choice from those who prefer paying by cash.[32] In fact, a new business model is emerging to provide ad-free services to users who are willing to pay a fair amount of subscription fee.[33]

(b) Consumer benefits

Apart from the argument that ad-supported free content is the preferred option, the marketing sector is also trying to prove that Internet users are actually better off with OBA, whether they like it or not. It is a common focus of industry reports to quantify the benefits brought about by the online advertising ecosystem. In 2010, McKinsey & Company undertook an independent assessment commissioned by IAB Europe to evaluate the

---

[32] That said, allowing individuals users to make their own decisions might lead to a divide between 'data haves' and 'data have-nots', and might run counter to the 'collective good' nature of privacy, but measures to minimise these risks are not part of this section.

[33] Mimi An, *Why People Block Ads: And What It Means for Marketers and Advertisers* (2016) 17-20 <https://cdn2.hubspot.net/hubfs/53/assets/hubspot.com/research/reports/Why_People_Block_Ads.pdf.zip> accessed 16 March 2017. However, it should be noted not all of these services have made clear whether they simply stop displaying ads to subscribers, or also stop tracking subscribers as well.

consumer surplus of online services funded by advertising.[34] Based on the willing-to-pay level, it is estimated that an average household benefits as much as €38 per month from online advertising.[35] A similar, yet lower figure concerning the UK was cited in a 2015 Advertising Association report, in which the value created by ad-funded search, email and social media services amounts roughly to £90 per person per year.[36] Despite the gaps between the exact numbers evaluated in different studies, it is nearly indisputable that all individual Internet users have more or less benefited from free services and content online, which are largely financed by the online advertising business. If it is true that OBA can actually boost the performance of advertising, then Internet users may — at least arguably — gain a greater part of consumer surplus in terms of the amount and quality of these free goods. From that point of view, the industry's argument that the public have now better access to digital commodities does have at least some merits.

Having said that, it would be a separate issue whether the actual benefits have been exaggerated, and whether individual users have been fairly compensated. A shared limitation that weakens the credibility of both studies consists in the reliance on the user evaluation, not the actual value, of online services. The figures calculated on a willing-to-pay basis may have been overestimated, in particular when the participants are unaware of the actual economic value of their data. Some service providers disclose their average revenue per user (ARPU), which should be a more accurate indicator of the value of an average user's data. Facebook is a good case in point in that, unlike

---

[34] IAB Europe, *Consumers Driving the Digital Uptake: The Economic Value of Online Advertising-Based Services for Consumers* (2010) <http://www.iabeurope.eu/files/7113/7000/0832/white_paper_consumers_driving_the_digital_uptake.pdf> accessed 20 February 2015.

[35] ibid 15.

[36] Advertising Association, *Advertising Pays 3: The Value of Advertising to the UK's Culture, Media and Sport* (2015) 37 <http://www.adassoc.org.uk/wp-content/uploads/2015/01/Advertising-Pays-3.pdf> accessed 20 February 2015.

Google, its service is narrow in scope (only social media) and the ARPU is clearly listed in their annual report. In 2014, Facebook's annual ARPU in Europe is $11.60 (approximately £8.90),[37] which is significantly lower than the figure of £28 for the same year as cited by the Advertising Association report.[38] There is thus clear evidence that Internet users tend to overestimate the amount of money they would have to pay were they charged for online services. It is very likely that, if given those exact numbers rather than asked to figure them out, a portion of those participants would switch to paid-for, ad-free services.

2.2 Good for commerce? Industrial interests and digital economy

2.2.1 OBA as an emerging industry

As demonstrated in Chapter 1, online advertising has overtaken or is overtaking television advertising as the most popular marketing channel in terms of budget in many countries. At the same time, OBA is becoming the mainstream in digital marketing with a strong growth. Compared with other forms of advertising, OBA is indeed a very young business, but it has been developing so fast that today, the industrial interests it represents are no longer negligible. There should be little doubt that the business model of the advertising industry, characterised by the advertiser-intermediary-publisher value chain, is in general legitimate. Indeed, even the GDPR has explicitly

[37] Facebook Inc., *Form 10-K* (2016) 37 <https://s21.q4cdn.com/399680738/files/doc_financials/annual_reports/2015-Annual-Report.pdf> accessed 22 February 2017. The latest figure for the year of 2016 is $19.42 (See Facebook Inc., *Facebook Annual Report 2016* (2017) 36 <https://s21.q4cdn.com/399680738/files/doc_financials/annual_reports/FB_AR_2016_FINAL.pdf> accessed 1 August 2017). Please note that the user base used for calculation here includes only monthly active users (MAU). If those who are less active are also accounted for, the ARPU outcome would be even lesser.

[38] Advertising Association (n 36) 37.

recognised direct marketing as a potential legitimate interest.[39] Since OBA forms an important part of direct marketing, it is reasonable to believe that its normal business deserves a certain degree of protection. Conceivably the OBA sector might strive for such protection from at least two aspects: first, it should not be subject to unreasonable regulatory constraints; second, it should be supported in the international competition on a level playing field.

(a) Deregulation

During the legislative process, the GDPR is sometimes criticised by the online advertising sector for imposing too strict a regime on the use of data. To them, the stricter rules introduced by the GDPR are doomed to stifle the promising future prospect for the industry. Soon after the European Commission published their very first draft of the GDPR proposal in 2012, trade groups in the marketing sector began to react critically. Trade organisations jointly signed an open letter to the UK government to express their concerns.[40] The proposal, they believed, would 'not just risk chilling the evolution of business models' but 'also place significant burdens on existing businesses', which jeopardises 'a policy and business environment that ensures low barriers to market entry, enables them to "scale up", and allows them to challenge established businesses.'[41] They were later joined by a European coalition of industry bodies, who urged the European Parliament to make sure this is

[39] 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] OJ L119/1 ('GDPR'), Recital 47.
[40] Sarah Shearman, 'Europe's "chilling" Data Reforms Will Deter a UK Facebook or Twitter' *Campaign* (2 May 2012) <http://www.campaignlive.co.uk/article/europes-chilling-data-reforms-will-deter-uk-facebook-twitter/1129829> accessed 24 March 2017.
[41] ibid.

'balanced with a need to allow for new business models and innovation.'[42] As the updated legal text proceeded through the legislative process, these bodies' statements turned even more strongly-worded, criticising the draft GDPR as a 'blunt-instrument approach'[43], 'setback'[44] and 'wrong turn'[45].

To the industry, the stringent legal framework would be unhelpful as the duties it imposes would eventually fall mostly on small- and medium-sized enterprises, throttling the possible growth of start-ups based on innovative use of data. Indeed, a more relaxed regulatory environment would mean fewer restrictions on how companies may monetise the data they hold. Before DoubleClick came into public attention and later under regulatory scrutiny in the US, it found a way to look into the massive online data and then pipeline the right advertisements to the right place. By successfully executing the idea of digging into data, it turned from a $50,000-turnover company with 12 employees to a multinational enterprise with 3,200 people and $500 million in revenues — in only four years.[46] Of course, from the industry's point of view, self-regulation or deregulation would be ideal as that means businesses in that sector may do as much as they want to maximise their profit. From a policymaking perspective, however, that does not stand on its own as a valid point. Just because reducing compliance costs may offer an industry the

---

[42] Internet Advertising Bureau UK, 'Industry Concerned over EU Data Protection Developments' (2013) <https://iabuk.net/news/industry-concerned-over-eu-data-protection-developments> accessed 24 March 2017.

[43] Internet Advertising Bureau UK, 'EU Governments Approve Privacy Text That Threatens Online Business' (2015) <https://iabuk.net/about/press/archive/eu-governments-approve-privacy-text-that-threatens-online-business> accessed 24 March 2017.

[44] IAB Europe, 'Data Protection Deal Is a Setback for Europe's Digital Economy' (2015) <https://www.iabeurope.eu/policy/data-protection-deal-is-a-setback-for-europes-digital-economy/> accessed 24 March 2017.

[45] Industry Coalition for Data Protection, 'Europe's New Data Rules Take a Wrong Turn' (2015) <http://epceurope.eu/wp-content/uploads/2015/12/ICDP-Press-Release_Final.pdf> accessed 24 March 2017.

[46] Mike Smith, *Targeted: How Technology Is Revolutionizing Advertising and the Way Companies Reach Consumers* (American Management Association 2015) 53.

opportunity to thrive does not automatically dismiss the necessity of regulation. It needs to be proved that the absence of regulatory measures would not lead to substantial risks to the economy, society, environment, and so on. Unfortunately, this does not seem to be the case considering all the threats potentially arising from the unfettered expansion of the OBA empire, as will be identified in next chapter.

(b) International competitiveness

The limitation-on-growth argument sometimes goes beyond the context of developing within an internal market, to a slightly different point that a too strict regulatory regime would undermine the competitiveness of the industry with a global view. For example, European Data Coalition, a lobbying group consisting of 21 European companies, warned in 2015 that the GDPR would 'cripple the EU outsourcing market and result in decreased levels of controller specialisation (as they will have to rely on in-house IT services) and ultimately a loss of competitiveness.'[47] An industrial leader expressly objected to the increasing regulation on the digital sector in Europe, calling it 'red tape' that would leave the EU 'lagging behind the US and in danger of being overtaken by China.'[48]

These claims also gain some support from the evidence collected by the EU policymakers. In 2012, a report had been prepared for the European Parliament's Committee on Industry, Research and Energy, in order to assess the regulatory impact on the EU's competiveness based on the then draft

[47] European Data Coalition, 'GDPR Redlines – The Missing Link Between GDPR and DSM' (2015) <http://europeandatacoalition.eu/wp-content/uploads/2015/06/GDPR-redlines-the-missing-link-between-the-dsm-and-the-gdpr.pdf> accessed 26 March 2017.
[48] Henry Foy, 'Google Warns Red Tape Threatens European Tech Sector ' *Financial Times* (17 April 2016) <https://www.ft.com/content/97185d98-0225-11e6-9cc4-27926f2b110c> accessed 26 March 2017.

GDPR.[49] Among the four sectors being assessed in the study (the other three being big data, cloud computing and privacy friendly technologies), behavioural advertising is considered as the sector whose competitiveness would be most heavily impacted.[50] The report picks up on a number of ways the draft GDPR might put EU advertising businesses at a disadvantage against their US competitors, including, for example, the provisions regarding consent:

> '[…] US based, globally operating, web based platform companies with massive user bases such as Google, Facebook, Amazon and eBay will be in a much better position to obtain consent. With strong B2C relationships, social reinforcement and critical mass acceptance, more frequent transactions covered by a single act of consent and important economies of scale they are more likely to achieve high consent rates than smaller companies and innovative start-ups, let alone predominantly B2B EU companies, who lack the end-user relationship required to achieve consent. The consent requirement may foster a more fragmented and closed EU internet where advanced targeting is dominated by US based platforms, instead of the open Digital Single Market envisaged in the Digital Agenda for Europe.'[51]

To sum up, what the OBA-related stakeholders argue against the EU data protection regime is that a lot of lucrative opportunities would shift to the other side of the Atlantic simply because the regulation there is more flexible and thus compliance costs are lower. However, several counter-arguments can also be made here. First, in cases where the investor targets the European market, any operation on EU consumer data would be equally subject to the EU's data protection legal framework, no matter where the business is incorporated. This has been made quite clear in the CJEU's judgment on the

---

[49] European Parliament, *Data Protection Review: Impact on EU Innovation and Competitiveness* (2012) <http://bookshop.europa.eu/is-bin/INTERSHOP.enfinity/WFS/EU-Bookshop-Site/en_GB/-/EUR/ViewPublication-Start?PublicationKey=BA3112305> accessed 26 March 2017.
[50] ibid 47.
[51] ibid 48.

high-profile case *Google Spain*[52], where the Court decided that Google, as a US-based data controller, was nevertheless liable to the duties under EU law. This is codified in a clearer way into the GDPR, which applies equally to data controllers established outside the EU, so long as the data processing in question relates to the offering of goods or services to EU data subjects.[53] That means, European and American competitors are in effect governed by the same set of rules when it comes to conducting business that targets the EU market.

Second, a high standard of data protection within an economy may help develop a healthy, reputable and high added-value industry. It is indeed very likely that in the short term, foreign clients may choose businesses in other jurisdictions to process non-EU personal data, so as to avoid the application of the GDPR. In the long run, however, this might turn into an advantage for the European industry as it gains reputation and experience in handling customer data with a world-leading standard. As an increasing number of consumers are becoming aware of the importance of online privacy, they are more likely to choose services that conform to stricter data protection rules.[54] This would create an incentive for international corporations to source data processing services from jurisdictions where personal data is believed to be properly safeguarded. Of course, this is at best only a possibility that is open to dispute. Still, this can stand as a counter-argument to the pessimistic belief that 'bad money drives out good' in the global digital market, which is not empirically verified, either.[55]

---

[52] Case C-131/12 *Google Spain and Google* [2014] OJ C 212/4.
[53] GDPR, art 3(2).
[54] Alessandro Mantelero, 'Competitive Value of Data Protection: The Impact of Data Protection Regulation on Online Behaviour' (2013) 3(4) International Data Privacy Law 229.
[55] For a detailed discussion of the 'race to the top or the bottom' debate surrounding international privacy standards, see Colin J. Bennett and Charles D. Raab, *The Governance of Privacy: Policy Instruments in Global Perspective* (The MIT Press 2006) ch 10; Graham

2.2.2 The wider digital market driven by data

The economic benefits of OBA, as the marketing sector often argues, are not limited to the industry itself but also extendible to a broader context of the digital economy. It is argued that online advertising forms an essential part of the driving force behind the growth of the digital market. An IAB-commissioned study conducted in 2012, for example, aimed to quantify the overall economic value generated by the advertising-supported Internet ecosystem.[56] The researchers employed three approaches to the evaluation of the Internet in the US: the boosted employment, the added output value and the time spent by users on the Internet. It is estimated that the 'employment due to the advertising-supported Internet ecosystem is 2.0 million direct jobs and 3.1 million indirect jobs, for a total employment of 5.1 million people.'[57] Besides, the Internet ecosystem is projected to have contributed $308 billion of direct value (ad-funded services, retail services and Internet access services) and $554 billion of indirect value.[58] They also converted the time of using the Internet into monetary value, which accounts for $760 billion, based on statistical data about Americans' usage of the Internet and average hourly wage.[59]

A similar assessment was carried out for the UK in 2013, commissioned by the Advertising Association to accounting firm Deloitte.[60] The study covered all categories of advertising in the UK, and as a general conclusion, it is

---

Greenleaf, *Asian Data Privacy Laws: Trade and Human Rights Perspectives* (Oxford University Press 2014) 559.

[56] Interactive Advertising Bureau, *Economic Value of the Advertising-Supported Internet Ecosystem* (2012) <http://www.iab.net/media/file/iab_Report_September-24-2012_4clr_v1.pdf> accessed 20 February 2015.

[57] ibid 80.

[58] ibid 81.

[59] ibid 81-82.

[60] Advertising Association, *Advertising Pays: How Advertising Fuels the UK Economy* (2013) <http://www.adassoc.org.uk/wp-content/uploads/2014/09/Advertising_Pays_Report.pdf> accessed 20 February 2015.

estimated that '[o]n average, £1 of advertising spend generates £6 for the economy.'[61] Given the high percentage of expenditures on online advertising (35.2% in 2013[62] and 47% in 2016[63]), applying that calculation to online advertising would roughly mean an economic contribution of £37.8 billion in 2013 and £59.7 billion in 2016. Some parts of the report have been written specifically with regard to online advertising. In terms of employment, for instance, it was estimated that digital advertising had created 9,317 jobs directly and 15,280 jobs indirectly.[64] The sheer benefit resulting from ad-funded search/referral services was valued at £2 billion and the increase on high street sales at £3.7 billion.[65]

It is however highly likely that the figures inferred in these studies do not reflect the precise value created by the online advertising sector. The suspicion comes not just from the motivation of these projects, which are mostly sponsored by organisations representing industrial interests, but also from the questionable methodology. For example, when calculating the employment boost of the online advertising sector, the IAB report attributes a total of the 2 million American jobs directly to the 'advertising-supported Internet ecosystem' without any evidence of the percentage of the Internet ecosystem that is actually funded by the advertising.[66] On top of that, a multiplier of 1.54 is applied to the calculation of indirect jobs created, again, without giving a

[61] ibid 9.
[62] Advertising Association and Warc, 'Shift to Digital Will Push UK Advertising to a Record £20bn+ in 2015' (2014) <http://expenditurereport.warc.com/FreeContent/Q4_2013.pdf> accessed 26 March 2017.
[63] Advertising Association and Warc, 'Adspend Growth Forecasts Hold Strong Despite Brexit Vote' (2017) <http://expenditurereport.warc.com/FreeContent/Q3_2016.pdf> accessed 26 March 2017.
[64] Advertising Association, *Advertising Pays* (n 60) 47.
[65] ibid 49.
[66]Interactive Advertising Bureau (n 56) 80.

clear explanation.[67] More importantly, it has not considered what if a large part of the Internet ecosystem turns to an alternative business model, or adopts alternative forms of online advertising. Assuming the entire ecosystem would not survive without advertising, or without personalised advertising, is ill-founded. Hence, the economic benefits of the OBA industry, as well as the costs caused by regulatory efforts, are actually highly challengeable. That said, this is not to deny *all* the economic significance of the sector — a large part of it should nevertheless be recognised. What is needed is perhaps a more trustworthy study based on accountable data, which is a matter of accuracy and objectivity.

2.3 Good for society? Innovation and democracy

2.3.1 (Ir)responsible innovation?

The spillover effect of OBA may be even stretched further to the common welfare of the society. This is usually articulated by the marketing sector with two desirable-sounding terms: innovation and democracy. Included in nearly all the industry reports and statements mentioned in the previous sections, one common concern is that an unbalanced data protection system would risk stifling innovation in the society. The logic behind appears to be that, first, innovation is always something valuable for the society, and second, innovation depends on the unconstrained use of personal data. Neither of the statements seems self-evident and therefore both would need to be examined more closely.

There is a growing body of research under the heading of 'responsible innovation' in science and technology studies (STS). One of the underlying

[67] ibid. The report has indeed cited a number of studies that have calculated the multiplier in similar sectors, ranging from 0.5 to 3.4, but has not explained how a number of 1.54 is eventually picked.

ideas is that '[o]nce we lift the lid on innovation to reveal its politics, we can start to see that, for all of the good intentions of individual researchers, innovation can be a form of what Ulrich Beck calls "organized irresponsibility."' [68] The potential threats to the environment or society stemming from innovative process should thus be taken into account as part of innovation policies. This would require innovation to be anticipatory of risks, reflective on practices, inclusive of public engagement and responsive to stakeholder values.[69] It should be pointed out that the theories of responsible innovation mainly address issues arising from *scientific* innovation, not *marketing* innovation or generally *commercial* innovation. In some areas of scientific innovation, such as stem cell[70] or nuclear energy[71] research, a novel breakthrough, no matter how promising, is often subject to constant debate over its ethical implications. That is because innovation is a social process that needs to accommodate public values.[72] As regards the innovative technologies that extensively make use of personal data, the ethical debate has however been disproportionately underwhelming. If even presumably well-intentioned scientific innovation might lead to undesirable consequences, there is no compelling reason to assume that innovation in the profit-seeking OBA sector is intrinsically desirable. Quite to the contrary, as will be further explored in the next chapter, business practices in this field often 'innovate'

---

[68] Jack Stilgoe, 'Foreword: Why Responsible Innovation?' in Richard Owen, John Bessant and Maggy Heintz (eds), *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society* (Wiley 2013) xii.

[69] Jack Stilgoe, Richard Owen and Phil Macnaghten, 'Developing a Framework for Responsible Innovation' (2013) 42 Research Policy 1568, 1570-1573.

[70] See Jeremy Sugarman and Douglas Sipp, 'Translational Stem Cell Research: Issues Beyond the Debate on the Moral Status of the Human Embryo' in Kristina Hug and Göran Hermerén (eds), *Stem Cell Biology and Regenerative Medicine* (Humana Press 2011).

[71] See Behnam Taebi and Sabine Roeser (eds), *The Ethics of Nuclear Energy: Risk, Justice, and Democracy in the post-Fukushima Era* (Cambridge University Press 2015).

[72] B Taebi and others, 'Responsible Innovation as an Endorsement of Public Values: The Need for Interdisciplinary Research ' (2014) 1 Journal of Responsible Innovation 118.

without sufficient assessment of impact, transparency or public participation. The innovation of the online advertising industry perhaps personifies the lack of responsibilities in innovative processes.

The second problem of the innovation argument lies in its oversimplified assumption that innovation would be suffocated without free use of personal data. Understanding why this statement is problematic would require an innovative perception of innovation. Cohen critically investigates the idea of innovation and its relationships with privacy. She points out that in the context of targeted advertising, profiling represents a form of knowledge, but a highly manipulative and powerful one.[73] She goes on to articulate why data privacy matters to innovation. A comparison is made between intellectual property law — which is intended to foster innovation — and the regulation over the use of personal data.[74] Both intellectual property and data protection regimes form some sort of institutional arrangements that ensure certain boundaries are respected.[75] Exactly as how allowing unauthorised use of copyrighted works would threaten a society's creativity, allowing unrestricted use of personal data would end up devastating a society's innovation, as the cornerstone of data sharing is trust. Of course, most if not all intellectual property laws also impose a list of exceptions to the right holders of copyrighted works and patents, and they are both subject to a maximum term. Similarly, the right to data protection is not an absolute one but often subject to restrictions. For example, under the GDPR, the use of personal data for archiving, research and statistical purpose is subject to a 'lighter' version of the purpose limitation principle.[76] However, in both cases, the scope of such

[73] Julie E. Cohen, 'Examined Lives: Informational Privacy and the Subject as Object' (2000) 52 Stanford Law Review 1373, 1405-1408.
[74] ibid 1416-1418.
[75] ibid.
[76] GDPR, art 5(1)(b).

exceptions should be limited to cases where there is an overriding interest, and the conditions for such exceptions should be firstly debated before the public, and then clearly defined by law.

Moreover, a true sense of innovation, as Cohen puts it, 'require[s] the ability to think outside or around existing, predictable technological and social pattern', which is incompatible with 'pervasive practices of monitoring, prediction, and preference-shaping'.[77] A practical implication of this insightful assertion could be that, a new source of innovation — and economic growth — may not rest with the old practice of digging into consumer behavioural data, but perhaps with the new paradigm of privacy-friendly economy instead. In the European Parliament report, despite the remarks that the draft GDPR might have negative impact on the behavioural advertising sector, it at the same time highlights a new area where the technology industry can benefit greatly from the legal reform: privacy-enhancing technologies (PETs).[78] It will remain to be seen whether the implementation of the new GDPR would actually open up such new opportunities for the European technology sector. Yet, it is quite clear that contradicting innovation with data protection does not provide a fully valid argument against legal restrictions over processing of personal data.

2.3.2 Ad-funded publishing and democracy

Last but hardly least, the online marketing industry occasionally makes reference to how OBA can be politically beneficial to the society as a whole. The starting point of this argument is that publishing matters to public debates in a democracy, and that advertising constitutes an important financial source of publishing. When it comes to a scenario of OBA, this argument is both old

[77] Cohen (n 73) 1427.
[78] European Parliament (n 49) 60-62.

and new. It is old because offline journalism has long been seeking advertisers as a major source of funds,[79] making it an easy argument to make. Yet, it also represents a new argument that the Internet, unlike traditional mass media, offers the audience with greater mobility.[80] Instead of reading passively from only a few local media in the pre-Internet era, people today are looking for news on a piece-by-piece basis from multiple sources in a much more active way.[81] Publications are thus said to have been 'unbundled' now, exposing people to opposing points of view.[82]

However, as will be further analysed in the next chapter regarding the implications of OBA on democracy, it is not difficult to explain why this is not the case. News readers today indeed have a much wider choice of what to see and hear. Yet, their time and attention remain scarce, leaving them only a limited capacity to receive and digest information. That means it is the filtering mechanisms that would determine what one actually sees, whether the filters are the user's choice or the algorithm's pick. This creates a new form of bundling, one that is not tied to particular media brands, but to the categorisation of readers, in a subtle, often indiscernible manner. It will be further illustrated how the filter bubbles as well as the consequent polarisation of society, both worsened by the pervasive practice of OBA profiling, would pose a serious threat to the democracy.

---

[79] Advertising Association, *Advertising Pays 3* (n 36) 40-43.

[80] Internet Advertising Bureau UK, *The Data Deal: How Data Driven Digital Advertising Benefits UK Citizens* (2014) 11 <http://www.iabuk.net/sites/default/files/The%20Data%20Deal%20-%20How%20Data%20Driven%20Digital%20Advertising%20Benefits%20UK%20Citizens.pdf> accessed 20 February 2015.

[81] Tow Center for Digital Journalism, *Post-Industrial Journalism: Adapting to the Present* (2012) 8 <http://towcenter.org/wp-content/uploads/2012/11/TOWCenter-Post_Industrial_Journalism.pdf> accessed 27 March 2017.

[82] Internet Advertising Bureau UK, *The Data Deal: How Data Driven Digital Advertising Benefits UK Citizens* (n 80) 11.

2.4 Summary: Validity and reality of the high expectations of big data

Having gone through the popular claims that are often made in favour of the OBA industry, we can now summarise those arguments and identify the valid ones. From an individual point of view, the 'relevant ads' claim does not seem capable of supporting the industry's position, as a user's choice of tailored or non-tailored ads should be respected. Free content and services may count as a valid argument at least partly, but it suffers from the exaggerated value of such 'free' products and also the inadequacy of an opt-out mechanism. As regards the trade arguments, they are not all convincing, either. The call for a lighter regulatory regime fails to articulate why the OBA practices are risk-free or at least low-risk enough for self-regulation; a high standard of data protection might be detrimental to the EU's online marketing businesses in the immediate future, but probably beneficial within a longer timeframe; the overall economy has actually benefited from the expanding OBA sector but again, the calculation might be biased. The slogan-like claim of 'good for innovation' holds very limited merits as innovation *per se* is not something intrinsically desirable. Neither can 'democracy' stand as a valid point because its advantage of diversifying sources of information is largely cancelled out by the concentration effect on individual or community level.

Therefore, if we are to name the 'legitimate interests' that are related to the use of big data in an OBA context, the ones remaining on the list would be the 'free content and services' and 'digital economy' arguments.[83] However, it should be noted that, despite the same terminology, the 'legitimate interests' identified in this chapter are not necessarily the same as those weighed up in

[83] See Guido Noto La Diega, 'Some Considerations on Intelligent Online Behavioural Advertising' (2018) 66-67 Revue du droit des technologies de l'information 53, 86-87.

the GDPR's balancing test. It will be pointed out in Chapter 4 that one of the legal bases for lawful processing of personal data is the necessity for the legitimate interests of the data controller or a third party.[84] Yet, according to the interpretation by the Article 29 Working Party, such interests must be lawful, sufficiently specific, and real and present in nature.[85] While 'direct marketing' is expressly recognised by the GDPR as a legitimate interest,[86] it is unclear how 'free content and services' or 'boosting the digital economy' can be specific enough as the *controller* or a *third party*'s real and present interest.

In addition, the fact, if proved, that OBA indeed represents a number of valid, legitimate interests does not mean that these interests can be pursued at all costs. Rather, the practices of intensive data use need to be justified on the balanced contemplation taking into account not only the benefits, but also the risks on the other side of the scale. An in-depth analysis of the values that the pervasive OBA-related activities might put in danger is therefore necessary, and that will be the main task undertaken by the next chapter.

[84] GDPR, art 6(1)(f).
[85] 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' (2014) 844/14/EN WP 217, 25.
[86] GDPR, Recital 47.

Chapter 3 Big Challenges of Big Data: A Theory of Big Data Risks in the Context of Online Behavioural Advertising

The big corporations that are dominating the Internet have recently been the subject of fierce controversy. Some are acclaimed as 'the most important contributors to the revolution of computers and technology'[1] while at the same time discredited as untrustworthy.[2] Some are regarded as the hope for ordinary people to take on repressive regimes[3] while at the same time condemned as a threat to democracy.[4] Those controversies hold true for the case of online advertising as well. As highlighted in the last chapter, trade associations representing the industry often claim that the entire ecosystem of the Internet has been supported by the advertising industry.[5] They also claim that without the constant financial stream coming from the advertisers, free services, news, videos and apps will all vanish.[6] After all, marketing is not something new at all. Long before the boom of the Internet, commercials on television, radio or newspaper had already been accepted as a fundamental financial source for these content providers.

---

[1] Steven Levy, *In the Plex: How Google Thinks, Works, and Shapes Our Lives* (Simon & Schuster 2011) 3.
[2] Scott Cleland, *Search & Destroy: Why You Can't Trust Google Inc.* (Telescope Books 2011).
[3] David Kirkpatrick, *The Facebook Effect: The Inside Story of the Company That Is Connecting the World* (Simon & Schuster 2010) 8.
[4] José Marichal, *Facebook Democracy: The Architecture of Disclosure and the Threat to Public Life* (Routledge 2016).
[5] IAB Europe, *Consumers Driving the Digital Uptake: The Economic Value of Online Advertising-Based Services for Consumers* (2010) <http://www.iabeurope.eu/files/7113/7000/0832/white_paper_consumers_driving_the_digital_uptake.pdf> accessed 20 February 2015.
[6] Internet Advertising Bureau UK, *The Data Deal: How Data Driven Digital Advertising Benefits UK Citizens* (2014) <http://www.iabuk.net/sites/default/files/The%20Data%20Deal%20-%20How%20Data%20Driven%20Digital%20Advertising%20Benefits%20UK%20Citizens.pdf> accessed 20 February 2015.

Having a relatively clear picture of the reality within the OBA industry as outlined in Chapter 1, one can easily point out the difference between the traditional forms of advertising and the ones presented on the Internet. While commercial messages on pre-Internet media might be readily tailored to their target audience, it can hardly be done on an individual scale. The revolutionary architecture of the Internet has enabled them to decide what and to whom to showcase their products or services on a user-by-user basis. A newspaper that can read the reader's mind might only exist in fantasy or sci-fi literature (so far), but in the world of the Internet, this is no magic at all. All it takes to personalise the content is simply a small piece of code on the webpage as well as a standard web browser on the users' side.

Such facts do not automatically place these companies on the evil side. They are not necessarily insincere when they argue that all they do is simply 'to get to know their customers better, and to provide them with more directly tailored services.'[7] In fact, as shown above, the whole point of constructing such a complex web of technologies is just 'getting the right ad in front of you.'[8] As one commentator puts it, Google is neither evil nor morally good;[9] at the end of the day, it is essentially just an advertising company.[10] However, just because they do not mean to be immoral does not mean that they are indeed benign. It might well be an authentic statement that they do not wish to do

[7] Sergey Filippov, *Data-Driven Business Models: Powering Startups in the Digital Age* (2014) 18-19 <www.lisboncouncil.net/component/downloads/?id=1081> accessed 1 November 2016.
[8] Craig Dempster and John Lee, *The Rise of the Platform Marketer: Performance Marketing with Google, Facebook, and Twitter, Plus the Latest High-Growth Digital Advertising Platforms* (John Wiley & Sons, Inc. 2015) 23.
[9] Siva Vaidhyanathan, *The Googlization of Everything (and Why We Should Worry)* (University of California Press 2011) 4.
[10] In the financial reports of Alphabet (Google's parent company), it is self-identified as a corporation that 'generate[s] revenue primarily by delivering relevant, cost-effective online advertising.' See Alphabet Inc., *Form 10-K* (2018) 53 <https://abc.xyz/investor/pdf/20171231_alphabet_10K.pdf> accessed 8 June 2018.

harms — which is not entirely relevant when it is what they do, not what they think, that matters — or even that they are actually contributing to the society — which has been challenged in the previous chapter. Nevertheless, these do not cancel out the side effects resulting from their practices. The debate about the ethics of OBA should not be framed as the choice of label between 'creepy voyeur' and 'humble servant'. Rather, it is about both the positive and negative effects of the increasingly pervasive presence of online tracking, profiling and targeting, in particular the less immediate, tangible or individualistic ones. The rest of this chapter will therefore seek to identify and clarify the genuine values at stake, and in the last section, to explore an alternative approach to the theorisation of these risks.

3.1 Private sphere facing invasion

3.1.1 Privacy as a fundamental right

Privacy is perhaps an obvious start of the exploration of big data risks. Many critics take a privacy approach to theoretic accounts for the dangers of the wide-spread application of big data, either in general terms[11] or in particular contexts like advertising.[12] Since privacy has been widely accepted as something both morally and legally desirable, campaigns that promote privacy protection may conceivably have greater appeal to the general public.

[11] See, for example, Terence Craig and Mary E. Ludloff, *Privacy and Big Data* (O'Reilly 2011); Julie E. Cohen, 'What Privacy Is For' (2013) 126 Harvard Law Review 1904; Kate Crawford and Jason Schultz, 'Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms' (2014) 55 Boston College Law Review 93; Ira S. Rubinstein, 'Big Data: The End of Privacy or a New Beginning?' (2013) 3(2) International Data Privacy Law 74; Omer Tene and Jules Polonetsky, 'Big Data for All: Privacy and User Control in the Age of Analytics' (2013) 11(5) Northwestern Journal of Technology and Intellectual Property 239.
[12] See, for example, Frederik J. Zuiderveen Borgesius, *Improving Privacy Protection in the Area of Behavioural Targeting* (Kluwer Law International 2015); Steven C. Bennett, 'Regulating Online Behavioral Advertising' (2011) 44 The John Marshall Law Review 899; Avi Goldfarb and Catherine E. Tucker, 'Privacy Regulation and Online Advertising' (2011) 57(1) Management Science 57.

For these reasons, when talking about the concerns with regard to OBA, 'privacy' would probably come first across one's mind. Much less straightforward and consensually acknowledged, however, is the essence and merits of privacy.

In Europe, privacy is one of the fundamental rights that constitutional laws guarantee. The European Convention on Human Rights (ECHR) does not adopt the exact wording of 'privacy' but affords everyone an express protection of the 'right to respect for his private and family life, his home and his correspondence' in Article 8.[13] Likewise, Article 7 of the Charter of Fundamental Rights of the European Union protects the 'right to respect for his or her private and family life, home and communications'.[14] At national level, many European countries' constitutions have also incorporated the protection of privacy (or 'private and family life') as a fundamental right.[15] The most intriguing part of looking into privacy issues from a European perspective is that, in the EU legal order, privacy and data protection are treated as two separate, albeit often interrelated, matters. This is evidenced by the fact that the Charter stipulates the right to privacy in Article 7 and the right to data protection in Article 8.[16] The change of wording is also manifest in the

[13] European Convention on Human Rights ('ECHR'), art 8.

[14] Charter of Fundamental Rights of the European Union [2012] OJ C326/391 ('Charter'), art 7.

[15] These include Constitution of Belgium [2014], art 22; Constitution of Bulgaria [2007], art 32; Constitution of Croatia [2010], art 35; Constitution of Estonia [2011], art 26; Constitution of Finland [2011], s 10; Constitution of Greece [2008], art 9; Constitution of Hungary [2011], art VI; Constitution of Latvia [2014], art 96; Constitution of Lithuania [2006], art 22; Constitution of the Netherlands [2008], art 10; Constitution of Poland [2009], art 47; Constitution of Portugal [2005], art 26(1); Constitution of Romania [2003], art 26; Constitution of Slovakia [2014], art 19; Constitution of Spain [2011], s 18. The English-language versions of these constitutions are available on Constitute (https://www.constituteproject.org/?lang=en). See also Blanca R. Ruiz, *Privacy in Telecommunications: A European and an American Approach* (Kluwer Law International 1997) 22.

[16] Art 7 ('Respect for private and family life') reads: 'Everyone has the right to respect for his or her private and family life, home and communications.' Art 8 ('Protection of personal

secondary legislation. While Article 1(1) of the Data Protection Directive (DPD) states its object as protecting in particular the 'right to privacy with respect to the processing of personal data',[17] Article 1(2) of the General Data Protection Regulation (GDPR) has changed the expression to the 'right to the protection of personal data'.[18]

Those who are less familiar with data protection in an EU law context might find it difficult to understand the interactions between privacy and data protection. Indeed, an inquiry into the history of the emergence of data protection as a fundamental right would reveal that, the creation of the 'data protection' label, as well as its attachment to (and subsequent detachment from) the 'privacy' notion, originates at best artificially, and at worst as a result of sheer legal miscommunication.[19] The 1970s saw a wave of national legislation (including constitutional recognition) regulating automated processing of data in Europe, but without clear reference to the objective of protecting individual 'privacy'.[20] At the same time, a parallel thread of developments was also taking shape in the light of international cooperation. In the course of drafting international instruments concerning data processing and free flow of data,[21] the term 'data protection' was adopted and given

---

data') reads: '1. Everyone has the right to the protection of personal data concerning him or her. …'

[17] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281/31 ('DPD'), art 1(1).

[18] 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] OJ L119/1 ('GDPR'), art 1(2).

[19] Gloria González Fuster, *The Emergence of Personal Data Protection as a Fundamental Right of the EU* (Springer 2014) 254-257.

[20] ibid ch 3.

[21] In particular, see OECD Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data [1980] ('OECD Guidelines'); Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data [1981] ETS No. 108.

sufficient substance — i.e. rules governing automated processing of personal data — to become a separate sector of law on its own.[22] However, the language and objective of these instruments had been greatly influenced by the approach taken by the US. There, following Westin's classical re-definition of privacy as the ability to determine the communications of information about oneself, [23] the emerging issues surrounding computerised processing of personally identifiable information had been resolved largely by expanding the concept of 'privacy' to cover what is known as 'informational privacy'.[24] The concurrence of 'data protection' and 'privacy' in such international documents has profound impact on subsequent legislation at both EU and national levels. The DPD and the e-Privacy Directive, for instance, have been crafted in a way that highlights 'privacy' as their major value orientation.[25]

The ambiguity, and often misuse, of the concept 'privacy' in varying contexts, such as 'respect for private life' in the ECHR, 'right to be let alone' in the fundamentalist American understanding and 'control over personal information' in the updated definition, has largely led to serious confusion regarding the relations between 'data protection' and 'privacy'. Today, we are seeing a reversal of the development in EU law: As mentioned above, both the Charter of Fundamental Rights and the GDPR have intentionally signified the disconnection from privacy. This is probably why the term 'data privacy' is sometimes preferred. It reflects both the traditional approach of right to privacy and the contemporary reality of a data-driven society.[26] Much ink has

[22] González Fuster (n 19) ch 4.
[23] Alan F. Westin, *Privacy and Freedom* (The Bodley Head 1967).
[24] ibid 70.
[25] This is evident from the introductory recitals and Articles 1 of both Directives, as well as the title of the e-Privacy Directive.
[26] For a brief discussion of the choice of wording between 'privacy', 'data privacy' and 'data protection', see Lee A. Bygrave, *Data Privacy Law: An International Perspective* (Oxford

been spilt on the similarities and disparities between privacy and data protection.[27] It is however not the main task of this section to investigate those interactions. Rather, this section is intended to clarify what 'privacy', in a narrow sense, means and why it matters in the context of OBA. Other related values can of course be discussed under the heading of 'privacy' with a broader perception, but that will be left to subsequent sections. Focusing on the core part of privacy would have the benefit of sharpening the arguments and making the contrast more visible. Therefore, the term 'privacy' is used here in a relatively strict sense to steer the discussion towards a more reflective direction.

3.1.2 Defining elements of a classical understanding of privacy

We cannot talk about privacy productively without at least briefly looking at its origin that predates the Internet. The earliest development of the right to privacy, at least in the Anglo-American sphere, is usually accredited to Warren and Brandeis, whose article *The Right to Privacy* has long been acclaimed as a ground-breaking work in the area.[28] In their eyes, privacy is a desirable remedy provided by the law for the invasion of private and domestic life by the press.[29] Such an invasion starts from the increasing journalist interest in gossip about people's private relations and ends in the mental pain and distress of those suffering from unwanted publicity.[30] At that time, the idea of

University Press 2014) 1-4; Graham Greenleaf, *Asian Data Privacy Laws: Trade and Human Rights Perspectives* (Oxford University Press 2014) 5.

[27] For example, see Juliane Kokott and Christoph Sobotta, 'The Distinction Between Privacy and Data Protection in the Jurisprudence of the CJEU and the ECtHR' (2013) 3(4) International Data Privacy Law 222; González Fuster (n 19) ch 2; Menno Mostert and others, 'From Privacy to Data Protection in the EU: Implications for Big Data Health Research' (2017) 25(1) European Journal of Health Law 43.

[28] Samuel D. Warren and Louis D. Brandeis, 'The Right to Privacy' (1890) IV(5) Harvard Law Review 193.

[29] ibid 195-196.

[30] ibid 196.

privacy was characterised by two points: the circulation of private information and the harm to one's personality. In short, privacy deserves legal protection because otherwise people's natural desire to retreat from public observation will not be satisfied. European jurisdictions have taken a different approach, which will be discussed later in this chapter.

Privacy theories, of course, have remarkably developed since then, and so have technologies. In the US, one example of such notable progress is the taxonomy of privacy torts. Prosser, in his milestone article published in 1960, conducts an extensive sweep of American court decisions with a privacy element.[31] He summarises four typical forms of privacy breaches: intrusion, public disclosure of private facts, false light in the public eye and appropriation.[32] Except for the first type, all other three bear a hallmark of Warren and Brandeis's concept of privacy — the circulation of information. Also, these categories all involve the (mis)representation of the plaintiff in the public. The first category, namely intrusion upon one's seclusion or solitude, or into private affairs, does not require the element of making information public. This clearly departs from the position of Warren and Brandeis, who view, for instance, eavesdropping as being governed by law of trespass, not law of privacy.[33] By contrast, under the classification of Prosser, even the illegal search of one's shopping bag in a store would fall within the first category of privacy invasion.[34] Therefore, the focus has obviously shifted from 'circulation' to 'observation' of information, although the line between public and private spaces remains intact in all four cases.

[31] William L. Prosser, 'Privacy' (1960) 48(3) California Law Review 383.
[32] ibid.
[33] Warren and Brandeis (n 28) 212.
[34] Prosser (n 31) 389.

Such a development in the concept of privacy can arguably be attributed to the reality that physical boundaries drawn by property law (land, home or mailing) can no longer afford sufficient protection. It becomes even more so in an age of information. The Internet has provided an unprecedented space for new forms of media to flourish, and for ordinary people to collect and share information at very low costs. If newspapers in the late 19th century are the only ones who can efficiently discover and disseminate anecdote stories, the ease of both manual and automated collection and processing of personal data today has empowered almost everyone to do that job.[35] Therefore, the privacy threats of big data are often theorised in the light of how intrusive processing of personal data can be. Tene and Polonetsky develop the concept of 'incremental effect', a piecemeal process through which the profiles pertaining to individuals could become more revealing as a result of the accumulation of personal data.[36] Ohm depicts a similar scenario with what he terms the 'accretion problem'. In the event that separate pieces of information from anonymised databases are linked altogether, he argues, they can possibly be unlocked if only one of those pieces is associated with one's real identities, ending up with an almighty, giant database-in-the-sky.[37]

Their concerns are not simply pulled out of thin air or pure imagination, but rather tested by real-life cases. In 2006, AOL released a list of 20 million search queries that were de-identified with only a random ID assigned to each searcher. Shortly after the publication, *The New York Times* managed to identify a woman numbered as 'User 4417749' in person, merely from a number of

[35] See, for example, Kim McNamara, 'The Paparazzi Industry and New Media: The Evolving Production and Consumption of Celebrity News and Gossip Websites' (2011) 14(5) International Journal of Cultural Studies 515.
[36] Tene and Polonetsky (n 11) 251-252.
[37] Paul Ohm, 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' (2010) 57 UCLA Law Review 1701, 1746-1748.

keywords like 'numb fingers' and '60 single men'.[38] Once equipped with massive datasets about individuals' details, and with technologies facilitating data mining, any state, business or even individual could easily have a grasp of people's jobs, leisure activities, preferred supermarkets or other information as sensitive as medical conditions.[39] To that extent, monitoring someone's online activities does not seem substantially different from paparazzi taking photos of a celebrity's home.

3.1.3 A technological middle ground: Difficulties of privacy conceptualisation for OBA

Despite the similarities, in the case of data processing through new technologies, there are still a few characteristics that are different from what was experienced a century ago. They include the level of human intervention and the expected scope of private sphere. These features are worth further discussion, as they may constitute a compelling reason to differentiate the present situation from the past.

In the case of OBA, user data is almost always collected and analysed in a fully automated manner. In the traditional idea of privacy, however, the rationale of providing such legal protection is based on the assumption that people care about what they look like to others. Here, 'others' can be only one person (in the case of wiretapping) or the unspecified public (in the case of media disclosure), but either way, it is human beings who are observing. Does that sense of privacy cover the monitoring by machines? Arguably not. It can

[38] Michael Barbaro and Tom Zeller Jr., 'A Face Is Exposed for AOL Searcher No. 4417749' *The New York Times* (9 August 2006) <http://www.nytimes.com/2006/08/09/technology/09aol.html> accessed 7 January 2016.

[39] A 2009 study finds that it is possible to infer one's sexual orientation from their Facebook friendships. See Carter Jernigan and Behram F.T. Mistree, 'Gaydar: Facebook Friendships Expose Sexual Orientation' (2009) 14 First Monday <http://firstmonday.org/ojs/index.php/fm/article/viewArticle/2611> accessed 7 January 2017.

be refuted by the statement that '[c]omputers don't breach privacy — people do!'[40] The idea behind this statement might be that, since privacy is threatened only when one is subject to human observation,[41] the processing by purely automated means should not be regarded as a threat to privacy. It is true that those who operate such profiling databases can always delve into the data of a specific user, but what if that is limited to a minimal level by technological, organisational and even legal restrictions? It is of course technically possible for marketing analysts to do exactly the same as what *The New York Times* has done on their enormous database, but they have little, if any, interest in tracking down one particular customer in person. It is true that certain forms of automated processing of personal data are of course more sensitive and sometimes require human oversight. As will be seen in the next chapter, automated decision-making is one of the subject matters of EU data protection law. However, it is not completely clear at this point whether targeted advertising should count as a decision that 'significantly affects' the user.[42]

Another problem of examining OBA through the traditional lens of privacy is the unclear territory of private space. The underlying justification for privacy is that one's will to stay away from public exposure should be respected. The law thus provides an alternative option that, unless one indicates otherwise, what he or she does, says, writes or otherwise expresses in a relatively exclusive place (home, changing room, or even conference room in some cases) or manner (sealing a letter in an envelope) should not be subject to monitoring or publicity. When it comes to the Internet, however, the physical separation between private and external sphere is highly blurred.

[40] Craig and Ludloff (n 11) 15.
[41] Daniel J. Solove, 'Privacy and Power: Computer Databases and Metaphors for Information Privacy' (2001) 53 Stanford Law Review 1393, 1418.
[42] GDPR, art 22.

There is no doubt that one's electronic device is their property, but when a user visits a website by connecting the device to a remote server via the Internet, how far can the user's 'private space' expand? From a purely technical point of view, this can be interpreted in completely different manners. Some might see the tracking of online footprints as observing one's private life, because how one's device is used is in most cases expected not to be shared with others. Others might however see the recording of online history by the websites themselves quite reasonable, because when Internet users visit the website, they have arguably placed themselves within the domain of that website, which cannot be seen as acting in a private space. Cookies are stored in the device, but the content is set and read according to the script sent from the website. To the extent that a web browser takes orders from both the website and the user, and that both sides possess a (but unequal) degree of control, can cookies and other forms of device fingerprinting be considered essentially a territory shared by both parties?

This is even more complicated given how many actors are engaged in the loading of even just one webpage. As briefly outlined in Chapter 1, apart from the websites that the user means to visit, there are other third-party intermediaries engaged in the tracking, profiling and targeting processes. Most of these operations are not carried out directly between these services, but rather through the user's browser. In technical terms, these communications are 'requested' by the browser as per the commands programmed by the website. Should they be deemed as part of one's 'private and family life' as defined in the ECHR or the Charter, or within the territory for one 'to be let alone' as defined by Warren and Brandeis? Again, there might be polarised ways to narrate the underlying mechanisms, which may lead to quite opposing perceptions towards the actual level of invasiveness. For

example, one might contend that, even with the concession that the communications between the user and the website are protected as private and confidential, the interference by the third-party trackers is actually enabled by both sides and thus should not be considered as private anymore. In theory, the most common form of third-party tracking — by means of external reference — cannot take place without authorisation from both the website and the browser. However, the counter-argument can be that, since average users in most cases have no knowledge of the existence of third-party tracking, or have insufficient skills to disable the tracking, or have concerns over functionalities being restricted, their expectation remains that the communications should be nevertheless protected as if they were private conversations.

Maybe there is no comparable instance of the private/public division in the offline world that can perfectly capture how exactly electronic communications work. That might be the reason why the classical notion of privacy can hardly come into play. When American judges and scholars are still debating whether the emission of smartphone geolocation data should be considered 'voluntarily disclosed public movements' and thus not protected by the Fourth Amendment,[43] the limitation of privacy as a legal concept becomes manifest. This also explains why, while the idea that one can be 'private in public' is largely rejected in the US,[44] this is readily accepted in Europe.[45] The borders between private and non-private are not that clear-cut anymore in the case of the Internet. The private/public dichotomy is no longer

[43] Monu Bedi, 'The Curious Case of Cell Phone Location Data: Fourth Amendment Doctrine Mash-up' (2016) 110(2) Northwestern University Law Review 507.
[44] Ronald J. Krotoszynski, Jr., *Privacy Revisited: A Global Perspective on the Right to Be Left Alone* (Oxford University Press 2016) 5.
[45] N.A. Moreham, 'Privacy in Public Places' (2006) 65(3) Cambridge Law Journal 606.

capable of explaining the structure of cyberspace, not least in the case of OBA.[46] A user's electronic device is their extended arm to reach others, but it is also the latter's extended arm to reach the user.

The indistinct frontiers between the parties of e-communications are a product of the unique infrastructure and standardisation of the Internet. The HTTP protocol and cookies standards have largely determined the direction of data flows and the readiness of such data on request.[47] The configurations of the Internet are therefore bidirectional but asymmetric. The use of cookies, for instance, has greatly enhanced user experience and made new forms of interactions possible. However, it has also subjected the device to the instructions of the website for their own good. From that point of view, such configurations are both enabling and limiting for individual users. The one-way metaphor of the Panopticon fails to explain the active part that the 'prisoner' might play in the instance of the Internet. Unlike a prisoner in the Panopticon, a user on the Internet has a degree of control over — and sometimes even benefits from — the observation by the watchers. Still, the possibility for individual users to exercise their choice is largely subject to technical, economic and social constraints. Technically, counter-tracking techniques are available but not always effective. On the one hand, disabled or cleared cookies can be bypassed with new tracking devices (including Flash cookies, mobile phone unique IDs or other forms of device fingerprinting that enable the so-called 'cookie respawning');[48] on the other, despite the increasing popularity of smarter software or plug-ins designed to block online tracking,

---

[46] See Daniel J. Solove, *Understanding Privacy* (Harvard University Press 2008) 24.
[47] See Paul M. Schwartz, 'Privacy and Democracy in Cyberspace' (1999) 52 Vanderbilt Law Review 1609, 1621.
[48] Omer Tene and Jules Polonetsky, 'To Track or "Do Not Track": Advancing Transparency and Individual Control in Online Behavioural Advertising' (2012) 13(1) Minnesota Journal of Law, Science & Technology 281, pt II.

these measures are facing an arms race from online marketers who have deployed anti-adblocking solutions.[49] Economically, disabling cookies or using anti-tracking software may mean excessive burdens on Internet users, including frustrating browsing experience, missing information or even outright denial of service.[50] Socially, certain services may form an important part of the digital social life of a large group of people,[51] for whom retreating from such services (as the only means to stop their tracking) is not a feasible option. If opting out from tracking continues to be unfamiliar, inconvenient and impractical, it is conceivable that the majority of the society might be conditioned to no longer care about taking alternative options.

Hence, although the structure of the Internet is designed for both ends of the network, the technical, economic and social reality has in effect skewed the balance in favour of the online trackers. These factors may have a decisive impact upon what people expect to be private or public, but the problem is that these factors tend to be protean in an online environment. The boundaries of walls and envelopes are easier to identify, but in the case of online tracking techniques, the line is relatively hard to draw. The reduction of human interference as well as the increase of the general public's awareness and the possibility to regain control may easily redraw the expected line of private sphere and lead to a different conclusion.

---

[49] Rishab Nithyanand and others, 'Adblocking and Counter-Blocking: A Slice of the Arms Race' (2016) <https://arxiv.org/abs/1605.05077v2> accessed 12 December 2016.

[50] Ashkan Soltani and others, 'Flash Cookies and Privacy' (2009) 4 <https://ssrn.com/abstract=1446862> accessed 7 January 2017.

[51] Gwenn Schurgin O'Keeffe, Kathleen Clarke-Pearson and Council on Communications and Media, 'Clinical Report—The Impact of Social Media on Children, Adolescents, and Families' (2011) 127(4) Pediatrics 800, 801.

3.2 Informational self-determination under threat

In spite of the etymological connection between 'privacy' and 'private', the concept of privacy as it is used today has gone beyond the scope of merely private and family life. This is in particular evident in Continental Europe. In a 1983 landmark decision, the German Federal Constitutional Court found that the freedom of individuals to decide on whether to engage in certain activities could be significantly inhibited if they can no longer ascertain who knows what about them and when.[52] A right to informational self-determination has therefore been confirmed, based on the general right of personality as guaranteed by the Basic Law and not limited to domestic domains but applicable to the public aspects of personal activities, such as participation in an assembly.[53] A similar idea also appeared in the Portuguese Constitution as early as in 1976, which explicitly prohibited processing of data not only limited to private life, but political convictions and religious faith as well.[54] Viewed from this angle, the scope of privacy will be substantially expanded.

In fact, long before the German judgment, there had been scholarly discussion about a new paradigm for privacy. Westin, for instance, argues in his book *Privacy and Freedom* that '[p]rivacy is the claim of individuals, groups or institutions to determine when, how and to what extent information about them is communicated to others.'[55] Fried also suggests that '[p]rivacy is not simply an absence of information about us in the minds of others; rather it is the control we have over information about ourselves.'[56] For him, the construction of physical separation — such as one's house — is but one means

[52] BVerfGE 65, 1 [1965] s II. 1. a).
[53] ibid.
[54] Decreto de Aprovação da Constituição da República Portuguesa [1976], art 35(3).
[55] Westin (n 23) 374.
[56] Charles Fried, 'Privacy' (1968) 77(3) The Yale Law Journal 475, 482.

of control over information.[57] The ability to have such control is vital to one's liberty because if the information exceeds the social context in which it is meant to be construed,[58] it might create inconvenience for that person and deter them from speaking something that is not morally wrong but unpopular or unconventional.[59]

However, the justification for the right to informational self-determination is not as straightforward as it might seem at first glance. Determining the dissemination of one's own data could arguably be seen as a way to manipulate the information about themselves, as Posner has demonstrated from an economic standpoint.[60] Manipulation is not necessarily something negative here. Posner believes that private information can be demanded by others as an asset to inform their decision-making.[61] Therefore, the rationale behind privacy law granting control over private information is not much different from granting property rights: Voluntary exchange may help achieve the optimal equilibrium.[62] To that extent, moralist and economic theories of informational self-determination do share certain common ground. If individuals have no control over their personal data, they would be more likely to try to keep information to themselves. While Fried would call it an unjust restriction upon people's liberty to freely speak, Posner would say it is a suboptimal condition that impedes the yielding and communications of valuable information.

---

[57] ibid 493.
[58] According to Nissenbaum's theory, the 'contextual integrity' is breached in this case. See Helen Nissenbaum, *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford University Press 2010) 140ff.
[59] Fried (n 56) 483-484.
[60] Richard A. Posner, 'The Right of Privacy' (1978) 12 Georgia Law Review 393.
[61] ibid 396.
[62] ibid 397.

Maybe the more significant disagreement between the two strands lies in whether and how the self-determination can be limited. Fundamentalist privacy theorists are more inclined to treat informational self-determination as something closer to an absolute right. For them, the ability to have such control is a precondition for the formation of one's unique 'self' out of the rather homogeneous social context. Lynskey argues, for example, that an individual's public persona may have multiple facets, the conflation of which could hinder their self-development.[63] Likewise, Rouvroy and Poullet have pointed out the paramount significance of informational self-determination as an absolute element of one's personality and the construction of their identity.[64] Cohen sees privacy as the breathing room for individuals' dynamic subjectivity to emerge.[65]

The ability of people to decide on how they are reflected in the external world by means of limiting information disclosure should therefore not simply be devalued to the disguise of their true selves. The diversity of social fabrics depends heavily on the resistance to being observed, assessed and then moderated to a monotonous pattern. To deny such a possibility is to deny human agency. Having said that, informational self-determination is not, and should not be, an absolute right. This holds true both from a utilitarian and an individualistic point of view. Posner contends that forced disclosure of private information is sometimes desirable if the nature of the information would lead to externalities or if transaction costs for voluntary acquisition are too high.[66]

---

[63] Orla Lynskey, 'Deconstructing Data Protection: The "Added-value" of a Right to Data Protection in the EU Legal Order' (2014) 63 International and Comparative Law Quarterly 569, 590.

[64] Antoinette Rouvroy and Yves Poullet, 'The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Importance of Privacy for Democracy' in Serge Gutwirth and others (eds), *Reinventing Data Protection?* (Springer 2009) 51.

[65] Cohen (n 11) 1927-1932.

[66] Posner (n 60) 397-401.

Solove also advocates a similarly pragmatic approach that privacy issues should be decided based on the balance of societal interests on both sides of the scale.[67] At the same time, looking from an individualistic viewpoint, the shaping of personal identity is not a product of merely isolated subjectivity. Social modelling and social constraints also play an important role in sustaining a stable society. Cohen, therefore, observes that the subjectivity should emerge 'gradually, in ways that are substantially constrained but not rigidly determined by social shaping.'[68]

There is little doubt that Internet users now have much less control regarding the use of data collected from them. In pre-OBA times, marketing research was done largely by means of surveys. Consumers could decide whether they would like to take part in any of these surveys and decide what answers to give for each question. By contrast, online behavioural data is collected automatically with only at best the safeguards of 'opt-out' or 'implied consent'. There is no easy way to adjust the data to what we would like it to be before it is sent out. Moreover, when the data is collected by advertisers, ad publishers or ad network providers, the fate of such data is completely out of the users' control. The whole idea of big data is all about maximising data collection and analysis, irrespective of how it is collected. That sense of lack of control runs throughout almost the entire process of OBA. In the tracking phase, for instance, only computer experts would have enough knowledge to monitor the data traffic so as to ascertain what kinds of data about the user are collected. As for the profiling phase, the highly sophisticated big data algorithms, combined with the lack of transparency, would make it extremely difficult for users to find out what profiles have been constructed and what

[67] Solove, *Understanding Privacy* (n 46) 50.
[68] Cohen (n 11) 1910.

judgments, predictions or automated decisions have been made based on those profiles.[69] Individuals might end up being subject to 'algocratic' systems where data is collected in a covert way and to inform certain forms of decision-making whose reasoning is opaque to humans.[70]

The less certain question is, should such form of deprivation of control be allowed as an exception to informational self-determination? The issues arising from OBA, or big data in a bigger picture, present a hard case for this question, whether within Posner or Cohen's accounts: From a solely economic standpoint, the value of online behavioural data and the transaction costs to acquire such data, as shown in Chapter 2, are hard to measure and are subject to intense debates. Equally controversial is, from a more complex, libertarian stance, whether the losing control over behavioural data will constrain people's liberty by imposing a chilling effect over the Internet. This would have profound implications for both human dignity and social democracy, which will be further investigated in Sections 3.3 and 3.5 respectively. As a conclusion of this section, it is sufficient to note that, whether examined with an individualistic or a collectivistic language, informational self-determination is an important value, but a rather instrumentalist one. There are scholarly discussions on whether privacy (or data protection) should be considered an instrumental value or an intrinsic one.[71] Many scholars have however accepted

[69] To be fair, some of the website have now allowed their users to download a copy of all your data. How these data are collected and processed however remains highly unclear. See Google, 'Download Your Data' <https://support.google.com/accounts/answer/3024190> accessed 7 January 2017; Facebook, 'How Can I Download My Information from Facebook?' <https://www.facebook.com/help/212802592074644> accessed 7 January 2017.

[70] John Danaher, 'The Threat of Algocracy: Reality, Resistance and Accommodation' (2016) 29(3) Philosophy & Technology 245.

[71] See, for example, Fried (n 56); James H. Moor, 'Towards a Theory of Privacy in the Information Age' (1997) 27(3) ACM SIGCAS Computers and Society 27; Herman T. Tavani, 'Privacy and Security' in Duncan Langford (ed), *Internet Ethics* (Palgrave Macmillan 2000).

that it has at least some instrumental facets.[72] When understood as the capacity to control personal data, the right to informational self-determination has a strong tendency towards an intermediate value.[73] Courts and scholars often resort to further justifications like dignity and democracy.[74] It clearly serves some values that are even more fundamental. The evaluation of the impact of OBA should therefore move forward to these underpinning values.

3.3 Dignity, liberty and equality

3.3.1 Human dignity and its two dimensions

If the value of the right to informational self-determination is not self-evident, there is a need to trace back to the ethical grounds on which such a right can be justified. Human dignity is the most significant one among these values. Its overarching importance is highlighted by the EU Charter of Fundamental Rights, which incorporates the respect and protection of human dignity as its very first article.[75] As with any other legal instruments that enshrine human dignity,[76] the Charter does not itself spell out a clear definition of human dignity. Other clauses of Chapter 1 of the Charter (titled 'Dignity') protect individuals from threats to life (Article 2), integrity of the person (Article 3), torture (Article 4), and slavery or forced labour (Article 5). They have a common element requiring that everyone be treated as human with respect, and not subject to inhumane treatment. This is also in line with Kant's moral

[72] See, for example, Bernd Carsten Stahl, 'The Impact of the UK Human Rights Act 1998 on Privacy Protection in the Workplace' in Ramesh Subramanian (ed), *Computer Security, Privacy and Politics: Current Issues, Challenges, and Solutions* (IRM Press 2008) 59; Rouvroy and Poullet (n 64) 50.
[73] Rouvroy and Poullet (n 64) 50.
[74] ibid.
[75] Charter, art 1.
[76] Such as the Universal Declaration of Human Rights (UDHR), the International Covenant on Civil and Political Rights (ICCPR), and the International Covenant on Economic, Social and Cultural Rights (ICESCR).

concepts. He distinguishes two forms of ends: dignity and price.[77] He notes that dignity is the end itself and cannot be replaced by something else.[78] Human dignity therefore requires to '[s]o act as to treat humanity, whether in your own person or in that of any other, in every case at the same time as an end, never as a means only.'[79]

The very immediate implication of such a requirement, thus, would be that no one should be used only as a means to achieve other ends. From this angle, some critics of big data are particularly concerned that the unfettered use of personal data would potentially reduce people from subjects to mere objects. Lyon, for instance, warns that the increasingly intensive use of personal data by computers might have the risk of degrading individuals to mere commodities and subjecting human values to mere efficiency.[80] Citron and Pasquale share a similar concern, showing how the present-day practices of data-driven scoring could turn individuals into ranked and rated objects.[81] Such a phenomenon is often illustrated in the context of credit scoring. For the financial sector, the use of big data credit scoring could probably save the costs for them, because they can figure out the patterns of potential default and most of their customers would fit into these patterns.[82] For those who match the pattern but are in fact underrated, this is unfair as they are not treated as they should have been, but instead simply ignored for the benefits of the company. This will be further discussed in Section 3.3.3 with reference to discrimination.

---

[77] Immanuel Kant, *Groundwork for the Metaphysics of Morals* (Thomas K. Abbott tr, Broadview 2005) 93.
[78] ibid.
[79] ibid 88 (emphasis omitted).
[80] David Lyon, *The Electronic Eye: The Rise of Surveillance Society* (University of Minnesota Press 1994) 109.
[81] Danielle Keats Citron and Frank Pasquale, 'The Scored Society: Due Process for Automated Predictions' (2014) 89 Washington Law Review 1, 3.
[82] Zuiderveen Borgesius (n 12) 43.

Here, we only need to point out that the origin of such a danger lies at the very disrespect of individuals as humans, to the contrary of the requirement of human dignity. Even if the creditworthiness predictions are not accurate for each and every applicant, lenders may still benefit greatly from a big data system as it would indeed be better than sheer arbitrary decisions.[83] The cost is that applicants are not evaluated in a way that people as individuals deserve. They are reducible to the sum of their data,[84] which is disrespectful because they are ignored only for profit-seeking, and also unfair because quantifiable data cannot capture the texture of one's life.[85] This criticism would remain valid even in a less commercial scenario. For instance, while collecting and analysing healthcare data for medical research might sound more acceptable, probably few patient would agree to receive treatment that is advised by a fully automated system designed to apply the findings in a way that would cut the costs for human expertise. Here, the true concerns are that, first, certain decisions may impact some vital interests of people, and second, such decisions are made in an uncaring manner that prioritises corporate interests over individual well-being.

In the context of OBA, a counter-argument has been made that it has very little impact on substantive rights or opportunities. Unlike creditworthiness check, insurance assessment, crime prediction or other application of big data, advertising is a much less sensitive sector when it comes to the direct effect on individuals. The denial of a loan or the raise of insurance premiums will have some immediate and tangible consequences, and therefore it makes sense to

---

[83] See Viktor Mayer-Schönberger and Kenneth Cukier, *Big Data: A Revolution That Will Transform How We Live, Work and Think* (John Murray 2013) 48; Zuiderveen Borgesius (n 12) 42.

[84] Julie E. Cohen, 'Examined Lives: Informational Privacy and the Subject as Object' (2000) 52 Stanford Law Review 1373, 1405.

[85] Solove, 'Privacy and Power' (n 41) 1425.

require a higher standard of caution. If the companies fail to take into account the possible inaccuracy of their automated decision-making process or fail to provide reasonable information, safeguards and remedies, they can hardly excuse themselves from the accusation that they do not treat their customers with full respect. In that case, human dignity is arguably violated as they show no care for the welfare of their customers. As for OBA, however, the improvement of targeting could be something beneficial to the users. The objective of an OBA system is never to disregard the uniqueness of each individual, but rather to create the possibility for the merchants to understand the consumers and meet their unique demands. Hypocritical as this statement might sound, the benefits of OBA should not simply be disregarded. However, it does not by any means suggest that online marketers have the right to conduct their business the way they are doing now. What it implies is simply that personalisation of online commercial messages can be a win-win situation if it is directed towards a fairer direction that benefits not just one side.

Online marketers often (knowingly dis)miss the point that, while it is one thing to admit the benefits of targeted advertising, it is another to say that they can do whatever they want to improve the targeting performance. It is not about whether it is desirable to tailor advertisements according to the users' interests; it is more about whether it is acceptable to achieve that sort of optimisation with the technologies they are using now. Recalling the seemingly contradictory survey results about the attitudes of Internet users presented in Chapter 2, such a distinction would be even more evident. The implication here from a human dignity perspective is that, the effects[86] are not the only thing that dignity concerns, it also concerns the vehicles to realise

[86] 'Effect' is a preferred term here compared to 'goal' and 'objective', because 'effect' does not imply any subjective pursuance, but only stresses that certain consequences follow as a result.

those effects. It is clear and plain that in most cases we cannot force another adult to do something, even if it is factually beneficial for them to do so.[87] It follows that it does not matter whether Internet users are actually better off with OBA; the deprivation of choice itself will suffice to constitute grave disrespect to human dignity. In other words, ends do not necessarily justify means.

The most intractable part of OBA, however, is that it falls in a grey area in terms of the extent to which one should have the final say on certain matters. While the parents' interference in the choice of a marriage partner, for example, is supposed to be a severe violation of their children's dignity, the enforcement of traffic safety rules (e.g. requiring drivers and passengers to fasten seatbelts) can hardly be said to have disrespectfully violated one's autonomy.[88] This is particularly true not just because it is advisable for the driver to obey the rules, but also because it benefits the society to do so. The case of OBA shares some similarities: As concluded in the previous chapter, OBA does represent certain individual, economic and societal benefits (albeit often exaggerated), which cannot simply dismissed altogether. As such, the dignity argument itself requires more clarification in the context of OBA. This cannot be done only with a general reference to human dignity. We need to turn to its two particular dimensions: liberty and equality.

3.3.2 Liberty

Social relationships provide a prism through which the light of human dignity scatters into a spectrum.[89] On the one end, we can observe the relationship

[87] Joseph Raz, *Between Authority and Interpretation: On the Theory of Law and Practical Reason* (Oxford University Press 2009) 138.
[88] ibid.
[89] See Ronald Dworkin, *Sovereign Virtue: The Theory and Practice of Equality* (Harvard University Press 2000) 128.

between an individual and society — known as liberty. On the other, the relationship between individuals themselves can be discussed under the heading of equality.

Lyon argues that liberty is a preferable term over privacy when talking about the totalitarian tendencies in a surveillance society.[90] In fact, a lot of discussion surrounding privacy has been oriented to the goal of protecting individuals from public intervention, especially taking into account the ubiquitous electronic surveillance at present. Although the exact meaning of liberty is not immutable, and sometimes overlapping with self-determination, it is by and large described as an individual's unabridged natural right to follow their own will.[91] In this context, the metaphors of the Big Brother and the Panopticon can be better understood here as one's free will being restrained under surveillance, or merely visible yet unverifiable inspection.[92] This affects not only the well-being of individuals, but also the maintenance of a democracy. As shown in the case of Big Brother, people have no freedom in any sense as what they say and do are all under the monitoring of the state, let alone to oversee or criticise the activities of the dictator. Prisoners in the Panopticon are also compelled to obey the norms as any disobedience can be observed and punished. The fact that someone is under constant inspection therefore has very strong suppressive effects. The observation assisted by technologies can be even more so because the effect can be imposed anytime and anywhere without all-day inspectors.

---

[90] Lyon (n 80) 13.

[91] Crawford and Schultz (n 11) 111.

[92] Lyon (n 80) 70-71. See also Alessandro Spina, 'Risk Regulation of Big Data: Has the Time Arrived for a Paradigm Shift in EU Data Protection Law?' (2014) 2 European Journal of Risk Regulation 248, 251.

Surveillance comes not only from totalitarian states, but also from present-day decentralised consumerism.[93] The term 'Little Brother'[94], or 'Little Brothers'[95], is coined as a variation of the Big Brother metaphor in the private sector. Their mundane practices are less likely to cause public attention as they are generally thought to have neither interest nor power in interfering individuals' liberty. However, just because they have not yet done so does not mean that they cannot or they do not want to. Some giant businesses are in fact quite comparable to certain nations, and sometimes even more ambitious.[96] Political power may not be their ultimate goal, but if that can be converted to cash streams, they do have the motivation to exert their influence. It is no secret at all that sometimes a business decision can be more effective than national legislation[97] or political campaigns.[98]

Another metaphor departing from Big Brother gains inspiration from Kafka's *The Trial*.[99] In this novel, the protagonist is arrested, prosecuted, convicted and eventually executed without being informed of what crime is involved. This fictional work reveals how individuals might suffer from a sense of powerlessness and helplessness with no control over what is happening to them.[100] *The Trial* differs from *1984* in that, as Solove observes, it

---

[93] Lyon (n 80) 78.
[94] Marsha Morrow McLaughlin and Suzanne Vaupel, 'Constitutional Right of Privacy and Investigative Consumer Reports: Little Brother Is Watching You' (1975) 2 Hastings Constitutional Law Quarterly 773.
[95] Wendy R. Leibowitz, 'Personal Privacy and High Tech: Little Brothers Are Watching You' *The National Law Journal* (7 April 1997).
[96] See Julian Assange, 'Assange: Google Is Not What It Seems' *Newsweek* (23 October 2014) <http://europe.newsweek.com/assange-google-not-what-it-seems-279447> accessed 7 January 2017.
[97] See Ethan Chiel, 'When Google Is More Powerful than the Government' *Fusion* (11 May 2016) <http://fusion.net/story/301142/google-bans-payday-loan-ads-adwords/> accessed 7 January 2017.
[98] See Zoe Kleinman, 'Facebook to Update Trending Topics' *BBC* (24 May 2016) <http://www.bbc.co.uk/news/technology-36367216> accessed 7 January 2017.
[99] Solove, 'Privacy and Power' (n 41) 1419ff.
[100] ibid 1421.

strikes at the heart of contemporary data processing — not to suppress individuality but to study and exploit it.[101] Also, '*The Trial* captures the sense of helplessness, frustration, and vulnerability one experiences when a large bureaucratic organization has control over a vast dossier of details about one's life.' [102] To that extent, big data might effectively represent a means of unaccountable decision-making, as the responsibility to give utmost consideration to individual cases can be easily shrugged off by fostering a 'computer says no' mentality.[103]

The case of OBA perhaps makes more sense with the metaphor of *The Trial*. The bureaucratic data systems created by online marketers have put individuals in an absurdist situation where the intensive use of their data on a day-to-day basis leaves them with no clue whether this will end up with benefits or detriments to them. This alternative metaphor has its limitations as well. For one thing, only because it is not part of the marketers' agenda to suppress individuals does not mean that it has no suppressive effects on them.[104] One of my colleagues once told me that she from time to time resisted the temptation to watch online documentaries on terrorist activities because she did not want to appear to show interest or sympathy to terrorism. Her concern is not necessarily about putting herself into troubles as a consequence of her 'suspicious' pattern being flagged up to the authority, but the mere fear or dislike to be associated with a certain kind of image is strong enough to deter her from certain behaviour. For another, in the case of *The Trial*, the

[101] ibid 1417.

[102] ibid 1421.

[103] 'Computer says no' is the catchphrase of a character in the British comedy *Little Britain*, in which a character often refuses the requests for assistance from customers after consulting the computer and responding with 'computer says no'.

[104] The dangerous effects of surveillance are acknowledged by Solove himself as well, but he argues that this is not the most central and pernicious problem of databases. See Solove, 'Privacy and Power' (n 41) 1418.

characters are headed passively towards an unpredictable yet probably predefined ending on which they have no say at all. As regards online marketing, the users similarly seem to have lost control over how their data is processed and who will be brought to them to make a pitch. However, technically speaking, they have the options to keep away from tracking (by disabling cookies) or to evade it (by using privacy enhancing technologies). There is indeed something they can do, but again, it would be inconvenient and costly, and not always effective. These options allow users to opt out in the first place but once the users accept (or tolerate) them, engagement in the process is impossible and there is probably only one way out: leaving the Internet entirely.

The impact of OBA on individuals' liberty should therefore be illustrated with reference to both the Big (Little) Brother(s) and *The Trial*. It constraints people's daily online activities in quite an indiscernible way and limits their autonomy in deciding how they would like to appear to others. The 'others' here can be either human beings or computers, but in slightly different ways. The need to protect individuals from human observation has been largely addressed in a narrow, traditional sense of privacy. Then why is it imperative to care about how we appear to computers? This is probably a mixed fear of the unknown. Indeed, the use of behavioural data for OBA purposes would unlikely give rise to direct effects of what can be considered discriminatory treatment. However, the data as such may serve non-OBA purposes as well, which may be much more sensitive and problematic. One case in point is the use of social media data for credit scoring,[105] something that might give OBA data holders some 'innovative' ideas. A lot of decisions today are made by

[105] Telis Demos and Deepa Seetharaman, 'Facebook Isn't So Good at Judging Your Credit After All' *The Wall Street Journal* (24 February 2016) <http://www.wsj.com/articles/lenders-drop-plans-to-judge-you-by-your-facebook-friends-1456309801> accessed 7 January 2017.

completely automated means based on one's data profile. These could be decisions related to employment, education, financial status, business offers, and so on. Internet users have no idea whether and how their profiles will inform these sorts of decision-making, so they would rather play it safe. If it is natural for a job applicant to feel concerned about the interviewers' first impression of them, then the desire to create a particular image in the database will be something quite understandable. The need to maintain one's public image stems largely from the fear of discrimination, and that is another form of big data risks that needs further investigation.

3.3.3 Equality

Apart from the protection against public interference, human dignity also requires everyone to be treated equally. Traditionally, it is anti-discrimination laws' task to tackle these issues. This is usually achieved by identifying certain categories of the vulnerable groups, and then prohibiting discrimination based on these categories. For instance, Article 21 of the Charter enumerates a list of such bases: sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation.[106] These labels are listed on a non-exhaustive basis,[107] but the recognition of protection against new forms of discrimination would generally take a lot more time and efforts.

Existing studies have identified two kinds of dangers arising from big data with regard to discrimination. On the one hand, big data may facilitate and reinforce existing discriminatory treatments.[108] Unjust discrimination becomes

[106] Charter, art 21.
[107] The actual wording of art 21(1) is '[a]ny discrimination based on any ground such as […] shall be prohibited.' The phrase 'any ground such as' makes it clear that the bases on which a practice might be held discriminatory are not limited to those that follow.
[108] Crawford and Schultz (n 11) 100.

easier as new ways of collection and prediction of sensitive data have developed. Also, big data makes discrimination less detectable by 'black-boxing' the potentially discriminatory decision-making. Either the original data-set or the algorithms can be contaminated by biased preconceptions,[109] but the complicated data processing involved may 'mask' such inappropriate elements and render it harder to trace the origins.[110] Also, even though no discrimination is intended, the message generated from unbiased input data with unbiased algorithm can nevertheless reinforce existing prejudice.[111] Suppose that a minor racial (or otherwise protected) group, for instance, is found less employable in a big data-powered study. It is proved that no racial information or other relevant indicators are included and the data processing is entirely objective. The conclusion can only be that they are actually less employable. However, if they are less employable only because they are the victims of previous stigmatisation of their group in history, then the widespread application of this study in the labour market will only serve to carry this stigma even further, and possibly undetectably.[112]

On the other hand, and even more importantly, big data might create new forms of discrimination. The power of big data consists in its ability to draw out certain correlations from a massive pool of data. These correlations do not have to be one-factored, and do not have to be associated with any sensitive factor under anti-discrimination law.[113] For instance, a conclusion that 'the

[109] Engin Bozdag, 'Bias in Algorithmic Filtering and Personalization' (2013) 15 Ethics and Information Technology 209.
[110] Tal Z. Zarsky, 'Understanding Discrimination in the Scored Society' (2014) 89 Washington Law Review 1375.
[111] ibid; Christopher A. Summers, Robert W. Smith and Rebecca Walker Reczek, 'An Audience of One: Behaviorally Targeted Ads as Implied Social Labels' (2016) 43 Journal of Consumer Research 156.
[112] Zarsky (n 110) 1399-1400.
[113] Jiahong Chen, 'The Dangers of Accuracy: Exploring the Other Side of the Data Quality Principle' (2018) 4(1) European Data Protection Law Review 36, 40-42.

residents in a particular area of the city are more likely to commit crime', or that 'the loyal consumers of a particular brand who also work in a particular industry are more likely to default on loans', can have adverse effects on those who match these patterns. A person might be treated unfairly if they happen to be a member of these groups by chance. The entire group might suffer from what Citron and Pasquale call the 'negative spiral'[114] and end up being 'segregated' from the society.[115] This effect has in fact found itself in certain controversial areas, such as insurance[116] and crime mapping[117].

The essence of anti-discrimination law thus needs re-clarification here. As Zarky notes, '[a]ntidiscrimination policy is not only about assuring equal treatment to equals, but also about assuring that specific differences among individuals should be ignored.'[118] In both the case of stigmatised protected groups and that of unspecified unprotected groups, big data could have very profound implications — probably in negative ways. This kind of repercussion is less visible in a context of OBA, because the mere selection and presentation of advertisements can be hardly seen as 'treatment'. Marketing flights from Edinburgh to Edinburgh-based users and those from London to London-based users can barely be considered discriminatory, although it might indeed limit a user's opportunities to be exposed to airlines from elsewhere. The necessity to require marketers to ignore the unfair differences

---

[114] Citron and Pasquale (n 81) 32-33.
[115] Zarsky (n 110) 1406.
[116] In 2010, car insurer Admiral was accused of charging customers born outside the UK a higher premium. See Rupert Jones, 'Which? Accuses Admiral of Driver Discrimination' *The Guardian* (16 March 2010) <https://www.theguardian.com/business/2010/mar/16/which-accuses-admiral-insurance-discrimination> accessed 7 January 2017.
[117] In 2000, pizza restaurant chain Domino's was sued in the US for refusing to deliver to certain parts of some neighbourhoods. See U.S. Department of Justice Office of Justice Programs and National Institute of Justice, *Privacy in the Information Age: A Guide for Sharing Crime Maps and Spatial Data* (2001) <https://www.ncjrs.gov/pdffiles1/nij/188739.pdf> accessed 7 January 2017.
[118] Zarsky (n 110) 1382.

between individuals can be further justified in two strands of discussion. One has very immediate and tangible consequences on individual users, namely price discrimination; the other is even more dangerous for the society, but will only manifest itself in the long run — the 'filter bubble' effect.

We shall start with price discrimination.[119] Online price discrimination began to receive public attention in 2000, when Amazon was suspected to have implemented price discrimination based on cookies.[120] Since the public outrage following media coverage, personalised pricing has become a rare practice,[121] but some less blatant forms remain, such as search result sorting.[122] Bad as discrimination might sound, 'price discrimination' is not necessarily a disapproval term in the discipline of economics. Quite the opposite, economists tend to prove that under certain constraints, price discrimination may produce an increase in overall social welfare.[123] Restaurants offering student discount might be welcome; a retail chain selling products at different prices in different stores might be understandable; a theme park charging people with disabilities a higher fee might be offensive. These are all price discrimination but not equally acceptable. The problem with OBA-based price discrimination is that the signals of market demands are detected but not

[119] This writing of this part has benefited a lot from the discussion with Julie Brill (US FTC Commissioner) in Amsterdam Privacy Conference 2015.
[120] Frederik Zuiderveen Borgesius and Joost Poort, 'Online Price Discrimination and EU Data Privacy Law' (2017) 40(3) Journal of Consumer Policy, 349.
[121] ibid. See also Arvind Narayanan, 'Online Price Discrimination: Conspicuous by Its Absence' (2013) <https://33bits.org/2013/01/08/online-price-discrimination-conspicuous-by-its-absence/> accessed 7 January 2017; Office of Fair Trade, *Online Targeting of Advertising and Prices: A Market Study* (2010) <http://webarchive.nationalarchives.gov.uk/20140402142426/http:/www.oft.gov.uk/shared_oft/business_leaflets/659703/OFT1231.pdf> accessed 7 January 2017.
[122] Zuiderveen Borgesius and Poort (n 120) 350-353.
[123] Akiva A. Miller, 'What Do We Worry About When We Worry About Price Discrimination? The Law and Ethics of Using Personal Information for Pricing' (2014) 19 Journal of Technology Law and Policy 41, 63; Thomas M. Lenard and Paul H. Rubin, 'In Defense of Data: Information and the Costs of Privacy' (2010) 2(1) Policy & Internet 149, 163.

given consciously by the consumers themselves. Also, the pricing strategies and the basis on which consumers are segmented are often invisible to Internet users. Transparency in information and possibility of signalling are both preconditions for positive price discrimination to function. In a classical price discrimination scenario, if a customer finds out that a supermarket in a different neighbourhood offers more favourable prices than the one in their own neighbourhood, they may decide which one to go depending on how much they value the price difference and the costs. This is the way they give out their market signal. In an online setting, however, they would never be able to tell for sure if they are correctly categorised by an OBA system when buying an air ticket. Also, they would not be able to tell whether the criteria of discrimination are fair: Is he offered a higher price just because they live in a wealthy area, or because they have shown greater interest?[124] Those who do not have the skills or resources to work around the pricing scheme might end up being 'second-class citizens' or the 'data have-nots'[125] in the society. The asymmetries of information will dismiss the justifications for online price discrimination practices.

Beyond price discrimination, another kind of adverse effect of big data emerges in a way similar to the two forms of segregation discussed in previous paragraphs. Imagine someone happens to click on an ad of a fishing kit, and is therefore profiled by an OBA system as someone interested in fishing. Later she finds (maybe unconsciously) herself surrounded by marketing information more or less related to fishing: fishing courses, fishing backpacks,

---

[124] For more examples of big data application of price discrimination, see Executive Office of the President of the United States, *Big Data and Differential Pricing* (2015) <https://www.whitehouse.gov/sites/default/files/docs/Big_Data_Report_Nonembargo_v2.pdf> accessed 4 October 2016.

[125] William Wresch, *Disconnected: Haves and Have-Nots in the Information Age* (Rutgers University Press 1996).

fictional books (with a fishing element) and flights (to destinations with fishing points). The more she is exposed to these ads, the more likely she clicks through and sees information, the more likely she is (labelled as) a fishing lover, and in turn, the more likely she is served with fishing-related ads. Rauhofer sees this practice as demand-side manipulation of the customer, 'who is thus reduced to the picture he presents based on previous purchases, website visited and services used.'[126] At this rate, individuals will be segregated, not necessarily from the society, but from the full range of possibilities and opportunities, in what Pariser calls their own 'filter bubble'[127] or Turow calls the 'comfort zone'[128]. Indeed, it is impossible to experience every possibility in our life and most of the time we need to make a choice, but when we make a certain choice (e.g. picking up fishing instead of a foreign language), all these possibilities are open to us and we are aware of the alternatives. As regards OBA, however, if the decision is made or strengthened by automated means, the space for other options will be compressed significantly. This might sound already too far from the traditional understanding of discrimination, but they follow a same path of the negative spiral.

3.4 Unfair imbalance of power

Whereas human dignity is an imperative attribute that allows us to actualise our 'selves' through social engagement, we are also involved in a diversity of social activities in different capacities. One of the most important capacities

---

[126] Judith Rauhofer, 'Round and Round the Garden? Big Data, Small Government and the Balance of Power in the Information Age' (2014) 2014/06 University of Edinburgh School of Law Research Paper Series 10 <http://ssrn.com/abstract=2389981> accessed 22 October 2014. See also Mireille Hildebrandt, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology* (Edward Elgar 2015) 68-69.

[127] Eli Pariser, *The Filter Bubble: What the Internet Is Hiding from You* (The Penguin Press 2011).

[128] Joseph Turow, *The Daily You: How the New Advertising Industry is Defining Your Identity and Your World* (Yale University Press 2012) 196.

today is acting as a consumer. The engagement in exchange activities reflects a significant facet of our lives in an economic sense. If consumers' personal data are collected in the course of provision of goods or services, they are protected as data subjects. In that event, data protection constitutes a cornerstone of consumer welfare.[129] There is however evidence suggesting that unfair data processing practices take the form of not just visible breaches of consumers' privacy and dignity but also of gaining an unjust transactional edge over consumers.[130] Economists have built different models to map the market distortions caused by information asymmetries. [131] The power imbalance resulting from excessive collection and use of personal data is notably highlighted in existing research.

In a previous section about the narrow sense of privacy, it has been pointed out that the technical structure of the Internet is bidirectional but asymmetric. Online content and services are delivered from the server to the user, which is visible to both sides, whereas behavioural data is recorded and transmitted in ways that are much less transparent to the user. The communications taking place each time a webpage is visited resemble an exchange of goods — 'free' information and services at the cost of allowing the access to information about

[129] European Data Protection Supervisor, 'Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy' (2014) Preliminary Opinion of the European Data Protection Supervisor, 31-32.

[130] Rauhofer (n 126) 8; Lynskey (n 63) 594; Solove, 'Privacy and Power' (n 41) 1450; Alessandro Mantelero, 'The Future of Consumer Data Protection in the E.U.: Re-thinking the "Notice and Consent" Paradigm in the New Era of Predictive Analytics' (2014) 30 Computer Law & Security Review 643, 650.

[131] For instance, see George A. Akerlof, 'The Market for "Lemons": Quality Uncertainty and the Market Mechanism' (1970) 84(3) The Quarterly Journal of Economics 488; Michael Spence, 'Job Market Signaling' (1973) 87(3) The Quarterly Journal of Economics 355; Michael Rothschild and Joseph Stiglitz, 'Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information' (1976) 90(4) The Quarterly Journal of Economics 629; Steven A. Sharpe, 'Asymmetric Information, Bank Lending, and Implicit Contracts: A Stylized Model of Customer Relationships' (1990) XLV(4) The Journal of Finance 1069.

what the user has done online. The problem here, however, is that the design of how information flows throughout the Internet is skewed towards the collection of data from the client side. When the user clicks a link or enters an address in a browser, a standard HTTP request will be constructed so as to contain quite a lot of information, such as the browser version,[132] the webpage from which the user navigates,[133] the IP address (which roughly indicates geolocation),[134] and most notoriously, the cookies. These details are sent by default and not easy to hide. That means, the Internet standards require users to disclose such information 'voluntarily' in the first place as they communicate with a website.

Upon receiving such a request, a server will respond by sending the content of a webpage back to the user, but not always. Sometimes the server just simply takes the payment (in the form of information) and gives nothing back in return. One might wonder who would visit a website that offers nothing, but this is happening all the time in a setting of online tracking. In 2015, Facebook was revealed to have kept track of even non-Facebook users on third-party websites.[135] By employing third-party cookies, Facebook managed to inject its codes to these websites. More striking is that, these websites are not necessarily ad publishers, or at least not on Facebook's ad network. What they did was simply put a 'Like' button on their webpages,[136]

---

[132] W3C, 'Header Field Definitions' (1999) sec 14.43 <http://www.w3.org/Protocols/rfc2616/rfc2616-sec14.html> accessed 7 January 2017.

[133] ibid sec 14.36.

[134] IETF, 'Internet Protocol' (1981) sec 2.3 <https://tools.ietf.org/html/rfc791> accessed 7 January 2017.

[135] Samuel Gibbs, 'Facebook "Tracks All Visitors, Breaching EU Law"' *The Guardian* (31 March 2015) <https://www.theguardian.com/technology/2015/mar/31/facebook-tracks-all-visitors-breaching-eu-law-report> accessed 7 January 2017.

[136] A 'Like' button is a link from a webpage to a particular social media service that allows the user to react to the content of the page so the user's contacts on social media know how they feel about the content. It is often accompanied by a 'Share' or 'Tweet' button by which a user can share that page to their social network. Both functionalities require the website

in the hope of increasing website traffic. However, this would allow Facebook to monitor user activities on these websites, something not necessary what these websites meant to do. This is how the Internet is standardised in favour of these service providers: Users give away the details of their online activities by default, and sometimes gain nothing back, or more ironically, just the display of a 'Like' button. Privacy settings or privacy enhancing tools may be of some help, but it is not always unobtrusive or effective.[137]

The imbalance between consumers and businesses stems not just from the configurations of the web, but also from information asymmetries. In the context of OBA, information asymmetries are manifest with the lack of knowledge about the conditions under which personal data is 'sold' to businesses. As with copyrighted works, the use of personal data often comes at a price. Unlike selling movies or music for cash, however, data subjects are usually paid by means of 'free' online content or services. A copyright agreement often imposes certain conditions on the user, such as the prohibition on sharing the work to third parties, but in the world of marketing, data sharing is ubiquitous. Mozilla began to develop an add-on called 'Lightbeam' for its web browser Firefox in 2011, which was designed to record and visualise how third-party websites are connected in the background for data collection (most notably by means of third-party cookies).[138] An Internet user, enabled by this tool, reported that when he visited 31 routine websites,

operator to embed a JavaScript on their webpage, which enables the social media service (like Facebook in this case) to generate the link and set up cookies on the user's device, even without the user clicking the buttons.

[137] Disabling cookies or opting out from targeted advertising may significantly impact user experience. See Jeff Sovern, 'Opting in, Opting out, or No Options at All: The Fight for Control of Personal Information' (1999) 74 Washington Law Review 1033.

[138] Mozilla, 'About Lightbeam' <https://www.mozilla.org/en-GB/lightbeam/about/> accessed 7 January 2017.

'the number of sites [he] was connected to was a staggering 117.'[139] Lightbeam also shows how seemingly unrelated websites are actually connected by one or more data collectors. These facts are largely unknown to average Internet users, who would probably have changed their mind had they been informed — and given a choice.[140] When they give consent to the collection and use of their data (sometimes 'by continuing to use the website'), they are unaware of the consequences because of the high costs to obtain such information. Certain services, like Netflix, would not even work if cookies are disabled.[141] What makes it worse is that, unlike those trackers who are 'trackable' by Lightbeam, a large part of data processing is simply undetectable. How data is shared, analysed, combined and reused is entirely black-boxed by a sophisticated big data system. The users' insufficient knowledge of how big data has made the harvest of personal data much easier and the aggregation of it much more valuable gives the business side an unfair advantage.

Another threat that exacerbates the already unjust power imbalance is the users' lack of bargaining power, which results partly from information asymmetries but has more to do with market dominance. As highlighted in Chapter 1, a number of ad network providers (such as Google and Facebook) are also dominant service providers in different sectors of the online market. As a result, Internet users are not left with many viable choices in these sectors. One can of course stop using Google or switch from Facebook to other services,

---

[139] Robert Wojtkiewicz, 'A Day in the Light: A 24-Hour Experiment with Mozilla's Lightbeam' (2013) <https://thebottomline.as.ucsb.edu/2013/11/a-day-in-the-light-a-24-hour-experiment-with-mozillas-lightbeam> accessed 7 January 2017.

[140] See Wiebke Thode, Joachim Griesbaum and Thomas Mandl, '"I would have never allowed it": User Perception of Third-party Tracking and Implications for Display Advertising' (Re:inventing Information Science in the Networked Society Proceedings of the 14th International Symposium on Information Science (ISI 2015), Zadar, 19-21 May 2015).

[141] Netflix, 'Netflix Error N8011' <https://help.netflix.com/en/node/71> accessed 7 January 2017.

but the costs will be so high that they are in effect forced to accept, as no realistic alternatives are available. The use of big data may add more fuel to the fire because those who possess more resources to harness data will profit a lot more from big data, and thus, the power tends to concentrate to a few oligarchies.[142] Also, on the individual side, Rauhofer has pointed out how the intrinsic power asymmetries may render the weaker parties more susceptible to giving up their data,[143] which in turn strengthens the already unfair status quo.[144] More importantly, these services are connected to the rest of the web in one way or another through the advertising network. That means, even if a user decides to retreat from a certain service, it does not mean that they will be completely out of touch with that service provider, since the latter, or its partner, is probably one of the ad network providers of other websites that they have visited. Perhaps the only effective bargaining chip is to cut off entirely from the Internet.

As a last point, we should also recall the manipulative power gained by the marketers as discussed in the previous section. It is worth some reiteration, and further exploration, of the repercussions here in the light of business-consumer relationships. Consumers are arguably placed in a weaker position as their preferences are somehow distorted. It is of course true that the whole point of advertising is to encourage potential customers to make the purchase, and the entire industry will vanish if advertisers are not allowed to have any influence on consumers in any sense. However, what makes OBA stand out from other forms of advertising on traditional media is that consumers are

---

[142] Mantelero (n 130) 650.

[143] Judith Rauhofer, 'One Step Forward, Two Steps Back? Critical Observations on the Proposed Reform of the EU Data Protection Framework' (2013) 2013/17 University of Edinburgh School of Law Research Paper Series 6 <http://ssrn.com/abstract=2260967> accessed 24 October 2014.

[144] Rauhofer, 'Round and Round the Garden?' (n 126) 8-9.

targeted on an individual basis in accordance to their online behaviour. By detecting the 'soft spot' in their hearts, OBA gives businesses an edge over consumers, an informational edge that empowers the supply-side to appeal better to those they target.[145] This is achieved through the intense collection and analysis of individual consumer behaviour, often without their awareness, and this is why it is more effective, and more invasive.

Throughout this section, the interests of Internet users, in the role of consumers, have been the major concerns. A few dimensions of OBA, as a representative of big data, have been examined with regard to how it changes the dynamics between businesses and consumers in a commercial world: infrastructural power, informational power, bargaining power and manipulative power. This is not the end of the list. If we treat individuals not just as consumers, but also members of a desirable democracy, then it might lead us to the discussion of another kind of power: political power.

3.5 Our undermined democratic society

The potential threats of big data are not limited only to those from individualistic standpoints. It also has certain implications for what we considered ethically good for the people in a community as a whole. Individuals should be perceived not merely as consumers but also as citizens, [146] who deserve adequate political participation in a democratic regime. The big data impact in societal terms is often underestimated as its effects are more long-term and intangible. Still, these matters deserve a close observation. As highlighted by many researchers, the values we cherish in our democracies could potentially become eroded without our awareness. For

[145] See European Data Protection Supervisor, 'Opinion 3/2018 on online manipulation and personal data' (2018), 9.
[146] Rauhofer, 'One Step Forward, Two Steps Back?' (n 143) 20-21.

instance, Cohen notes that autonomous, unmonitored choice is essential for the diversity of speech and behaviour.[147] She then goes on to explain how a community can benefit from this kind of diversity in various manners, including *inter alia* reasoned participation in governance, encouraged innovation in markets, reinforced stability of social fabric and a collectively defined role for technologies in the society.[148] Similar points are also made by Schwartz, who warns that once surveillance becomes the norm, cyberspace will no longer function as a forum for deliberative democracy.[149] When people are aware or suspicious of being monitored, they would probably feel constrained and become subject to, say, self-censorship as a consequence of the chilling effect. [150] The monitoring of OBA networks is no exception. Consumer data might be less directly associated with censorship, but it does facilitate potential surveillance attempts. The chilling effect can be even more powerful in particular when it is widely reported that intelligence services like the NSA or the GCHQ have access to records of personal online[151] and offline[152] traces. If Internet users have concerns with what they read and what they say being monitored, even just by an OBA system, they might refrain from certain activities, including the perfectly legitimate ones. Even worse, under the slogan of '[i]f you've got nothing to hide, you shouldn't worry about

[147] Cohen, 'Examined Lives' (n 84) 1923-1925.
[148] ibid 1926-1928.
[149] Schwartz (n 47) 1651.
[150] Solove, *Understanding Privacy* (n 46) 109; Paul Ohm, 'The Rise and Fall of Invasive ISP Surveillance' [2009] University of Illinois Law Review 1417, 1459.
[151] Glenn Greenwald, 'XKeyscore: NSA Tool Collects "Nearly Everything a User Does on the Internet"' *The Guardian* (31 July 2013) <https://www.theguardian.com/world/2013/jul/31/nsa-top-secret-program-online-data> accessed 7 January.
[152] Brian Wheeler, 'GCHQ Could "Grab" UK Shopping Data, Committee Told' *BBC* (10 December 2015) <http://www.bbc.co.uk/news/uk-politics-35060064> accessed 7 January 2017.

government surveillance'[153], future generations might be conditioned less resistant to surveillance use of online behavioural data.

The sensitivity of OBA becomes even more evident when political campaigns are now actually advertisable. With the revelations about Cambridge Analytica using social media data to help campaigners target voters,[154] it is clear that political advertising can be powerful, and also controversial. In fact, not just third-party companies like Cambridge Analytica are playing a role in political targeting; major platform providers are also actively involved.[155] Both Google[156] and Facebook[157] have their own advertising products that are specifically designed for political campaigners. It is anything but new today for these campaigners to employ targeting technologies to reach their voters more effectively.[158] In a political market, as with any other markets, talking to the 'right' customers is of paramount importance. However, the application of OBA to political marketing, as discussed regarding manipulative power in the preceding section, could give the campaigners an unfair competitive edge as they target individual voters. This can be done by predicting an individual's favourite products, sports, music, etc., and then personalising their messages with reference to the

[153] Daniel J. Solove, *Nothing to Hide: The False Tradeoff Between Privacy and Security* (Yale University Press 2011) 1.
[154] Carole Cadwalladr and Emma Graham-Harrison, 'Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach' *The Guardian* (17 March 2018) <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election> accessed 10 June 2018.
[155] For details about how Facebook, for example, has engaged in US election campaigning, see Daniel Kreiss, *Prototype Politics: Technology-intensive Campaigning and the Data of Democracy* (Oxford University Press 2016) 175-181.
[156] Google, 'Win the Moments That Win Elections' <https://www.google.com/ads/elections/> accessed 7 January 2017.
[157] Facebook, 'Facebook Elections' <https://www.facebook.com/business/a/politics-industry> accessed 27 June 2016
[158] Sasha Issenberg, 'How President Obama's Campaign Used Big Data to Rally Individual Voters' (2013) 118(1) MIT Technology Review 38.

appropriate brands, teams or artists.[159] They can also find out those who are more likely to be mobilised,[160] and then, in Facebook's words, 'lower [the] cost by optimising the delivery of adverts to the people most likely to take action.'[161] This could be an effective way to promote a campaign, but would probably make those who are already inclined even more committed, and those who are considered 'untargetworthy' more marginalised in political life.

Beyond the chilling effect and political manipulation, some scholars have provided alternative accounts for some remoter, less observable negative effects arising from unfair data processing practices for non-surveillance purposes. Sunstein, for example, uses the term 'cyber-polarisation' to describe how online media could shape our minds into certain groups, severely contrary to social pluralism.[162] Turow[163], Tene and Polonetsky[164] have applied this theory to the advertising industry. They argue that it is indispensable for a vibrant society to maintain a healthy balance between those media targeting the entire society and those tailored to certain communities.[165] The boom in tailored media in the past decades, however, has driven the society towards the heavily segment-making extreme.[166]

To summarise, three dimensions of democratic values might be compromised by the uchecked practices of OBA. First, the tracking of Internet users, if not conducted in a transparent way, will cast a shadow over the free

---

[159] David Sumpter, *Outnumbered: From Facebook and Google to Fake News and Filter-bubbles – The Algorithms That Control Our Lives* (Bloomsbury 2018) 53-54.
[160] Patrick O'Connor, 'Political Ads Take Targeting to the Next Level' *The Wall Street Journal* (14 July 2014) <http://www.wsj.com/articles/political-ads-take-targeting-to-the-next-level-1405381606> accessed 7 January 2017.
[161] Facebook, 'Facebook Elections' (n 157).
[162] Cass R. Sunstein, *Republic.com 2.0* (Princeton University Press 2007) 46-49. See also Marichal (n 4) 65-74.
[163] Turow (n 128).
[164] Tene and Polonetsky, 'Big Data for All' (n 11) 252.
[165] Turow (n 128) 192
[166] ibid.

use of the Internet. Second, the profiling of individual preferences will give unfair advantage to certain camps. Third, the targeting of voters, if not making users informed of the alternatives, will give rise to the polarisation of the society. All these negative consequences will greatly impair the proper functioning of a democratic society in the long run. OBA is a product of private interests, but it might end up with some serious societal impacts. From this point of view, the case study of OBA is not just a representative of big data in the private sector, but also touches on some part of its application in public affairs.

3.6 The falling walls of the data world

The preceding sections have so far introduced the major theories of the potential threats of big data. They then have been examined with the case study of OBA. Some of these arguments are more relevant to the OBA industry while others are less so, but generally OBA may be investigated as a useful representative of big data. The question then is: What can we learn from these perspectives? What are the valid arguments regarding big data in the context of OBA? What common ground do these arguments share? Can these arguments be applied to a wider backdrop of big data developments?

To answer all these questions, it is essential that we find an appropriate narrative to make sense of the ways in which the use of big data has intensified all the threats flagged up above, whether the individualistic, collective or societal ones, and whether quantitatively or qualitatively. With the help of the case study of OBA, at least four dimensions of transformation can be revealed and should be reiterated here: First, the fine line between private and public space has been technically smudged and then redefined. A user's terminal device now no longer serves as a shield of information, but rather a mere node in the gigantic, complex network for generating, communicating and

consuming behavioural data. Even in Europe, where protecting 'private in public' is more widely admitted, this would have profound implications as the shift of the boundary would make certain practices look less invasive, and thus harder to defend against. Second, the uncontrollable digital footprint has been fused with the controllable profile voluntarily created by the user. While some online services offer users the chance to, for example, refine their online profiles by adding, removing or editing certain details, or even indicate their advertisement preferences, it only provides a sense of control, not the genuine control of the invisible profile. While there are some 'dash board'-like functionalities in OBA networks to allow users to calibrate the behind-the-scenes creation of the hidden profiles, their roles are limited. The actors of the network do not have the incentives to provide such a possibility, and they may argue that it is technically impossible to do so given the heterogeneity of the industry structure. Third, and as a related point, the construction of advertising profiles has been led towards a multi-layered, multi-faceted, and mutually informative model. As pointed out in Chapter 1, a massive number of actors in the OBA network may respectively keep a record of a same user's online activities, with varying degrees of focus, breadth and precision. These records may maintain certain amounts of differences but have the potential to inform, strengthen and rectify each other. Fourth, the interpretation of an individual user's personal life has largely been based on the demographic and behavioural patterns of others. All these user profiles in the interconnected databases thus, in themselves, make up an unmeasurable network in which everyone's digital image(s) is susceptible to those of others.

All these new socio-technical realities point towards the need for a coherent account for the substantial, additional dangers of big data. A useful conceptual metaphor to capture the real dangers exposed from the four dimension above

could be that, we are now experiencing the falling walls in our highly datarised world — in all four senses.[167] The 'walls' represent the techno-economic divide between four pairs of domains that used to be separate but is now largely intermingled: private and public realms, manageable and unmanageable profiles, one data representation portrayed by an entity and another by a different entity, and one's data representation and the others'. The rest of this section will serve to test this metaphor by illustrating how it can better depict the escalating risks from all the perspectives covered in previous sections.

3.6.1 Privacy

We have noted that the essential concept of privacy draws a line that separate one's private life to their public one. The information within the private life circle should be protected against unauthorised observation and circulation. By controlling how much information can go through the sluice gate to the public domain, individuals are in effect managing their public presence, or in Posner's words, manipulating information about them.[168] The classical definition of privacy as the 'right to be let alone'[169] does not imply that individuals do not care about their public life. Rather, it is more likely to be the case that they want to be let alone once in a while because they know they will eventually need to return to the public life and therefore do not want private matters to impede their future engagement in public life.[170] Privacy, thus, may serve the purpose of maintaining a good (or at least not bad) image to the public side of life.

---

[167] For a similar metaphor of 'the wall is suddenly gone', see Woodrow Hartzog, *Privacy's Blueprint: The Battle to Control the Design of New Technologies* (Harvard University Press 2018) 151.
[168] Posner (n 60).
[169] Warren and Brandeis (n 28).
[170] Posner (n 60).

In an age of the Internet, apart from a good image in real life, those who wish to connect actively with others will also need to maintain their decent presence online. For instance, it is very common these days for corporations to make use of mass media, corporate media and social media to build up and maintain a corporate reputation.[171] Similarly, blogging and tweeting have become popular among academics as a means to increase professional visibility and public impact.[172] A large portion of Internet users are also building up their online profiles to impress their friends or a broader audience. Such forms of online presence management share a key factor with their offline versions: They have a heavy focus on reputation. In other words, it is about the strategies to give an impression appealing to others — family, friends, colleagues, employers, fans, customers or anyone who happens to show interest in that person. Maintaining a good name is not the only but a major purpose of managing one's offline and online presence.

However, the case of OBA, or a lot of other forms of big data application, is associated with reputation in a less immediate but more delicate way. As noted in Section 3.1, one of the major difficulties of applying traditional privacy theories to an instance of OBA is that human observation is substantially minimised. Simply put, if being observed by an insect on the wall is not a privacy problem,[173] why should being observed by a computerised system be? In *The Black Box Society*, Pasquale employs the concept of 'digital reputation' to demonstrate the threats we are facing today in 'an Era of

[171] Grahame Dowling and Warren Weeks, 'Measuring Media Corporate Reputations' in Sabrina Helm, Kerstin Liehr-Gobbers and Christopher Storck (eds), *Reputation Management* (2011) 113.
[172] Gill Kirkup, 'Academic Blogging: Academic Practice and Academic Identity' (2010) 8(1) London Review of Education 75.
[173] Solove, 'Privacy and Power' (n 41) 1418.

Runaway Data'.[174] By way of example, he examines the latest trends of making use of scoring technologies in finance[175] and health sectors[176]. The use of credit scores and medical records are of course known as the most notorious examples of potential tarnishing effect on one's reputation.

With the prevalence of OBA, the digital reputation might become increasingly difficult to locate and manage, and the underlying user profile might become conflated with the *private* one, the *visible* one, the ones held by *various profilers*, and even the ones of *others*. The walls between what is private and what is public have largely evaporated in the online world. The structure of the Internet makes it easier for marketers to build up user profiles but does not enable the users to manage such presence. If the creation of a technological artefact tips the balance of existing interest distribution in an unfair way, policy reconfiguration might prove necessary. An electronic device is the property of its owner, but it serves not just its owner, but also those remote online marketers. For this reason, computers or web browsers should now be seen as only one section of the channel between Internet users and their service providers. Even worse, as we have seen in Chapter 1, tracking, profiling and targeting are all done in a multi-levelled, decentralised manner. Advertisers, ad publishers, ad network providers, bidding agencies and other actors may each have a profile of a given user. These profiles support each other but maintain certain differences. To such an extent, big data has made data reputation outright unmanageable: If credit scoring is based solely on loan history, one can easily figure out how to avoid the adverse factors or what went wrong when something happens; but if it is based on a complex set of

[174] Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard University Press 2015) ch 2.
[175] ibid 22-25.
[176] ibid 27-30.

data that is connected to an uncertain scope of private activities, hidden profiles, external sources and other customers, it would be impossible for one to audit the scoring system.

3.6.2 Informational self-determination

Apart from privacy, the notion of the 'falling walls' may also help us reflect on other values that we have discussed. Informational self-determination, for example, has a lot to do with such a concept. While we tend to conceive 'information' as pieces of message that would disclose certain events of our life, what might be ignored is that the overall digital image should also count as a form of personal information and thus be subject to the principle of self-determination. At the time when records of personal data were kept isolated and static, having control over those separate pieces of information would suffice for the purpose of determining one's digital self. Now that one's digitalised image has been entrenched in a network and dynamically affected by so many factors, it is simply impossible to achieve the goal of developing one's informational self by only controlling individual data points. The informational walls in the offline world with which the distance between people is maintained do not stand in the online world anymore. Big data has brought integration of personal information to such a new level that asking individuals to decide the fate of their information each time they give out such information is no long effective, or fair. The building of one's public presence is much harder without such personal room surrounded by the walls. In other words, in the age of big data, informational self-determination should cover not just the 'in and out' of information, but more importantly, the border of their own data realm. So understood, the very idea of informational self-determination is actually facing more challenges, and thus requires a new paradigm to achieve.

3.6.3 Human dignity

As noted above, informational self-determination has a strong instrumentalist flavour. It serves certain underpinning values such as liberty and equality. It has been explained why people's liberty might be at stake with the metaphor of the 'Little Brother(s)'. The fact that Internet users have no idea when, how, by whom and what kinds of their personal data is being utilised would cause a psychological effect on them, who might in turn feel constrained when surfing the web. This effect partly stems from the potential, often unknown, consequences of the exploitation of their profiles, which will be further explored below. Another often neglected origin of the chilling effect can be outlined better with reference to the concept of the 'falling walls'. The problem here is that an impression is created, albeit only by computers. While it is true that being watched by an insect on the wall is not a problem for most people, being watched by a silent person at home might have a stronger impact upon one's behaviour. Even if that person will never tell anybody else about anything they have seen, their presence and observation will nevertheless create a sense of unease. The difference here is probably that a person is able to understand and judge on humans' behaviour, whereas insects are not. In the case of OBA, all the tracking, profiling and targeting processes are meant to make sense of individual Internet users' behaviour. Such kind of understanding do not have to have any negative consequence on individuals or to serve any surveillance purpose for the government. The mere fact that one is being monitored by something that would make sense of their speech or actions would suffice to cause self-consciousness and mental tensions. The fact that OBA data might be 'repurposed' for non-OBA uses and slip through the 'revolving door'[177] to the public sector would even add greater weights to

[177] Rauhofer, 'Round and Round the Garden?' (n 126) 8.

such stress. In a world of the falling data walls, the chilling effect might be even greater because of the stronger sense of human involvement: A user might feel even more constrained as the monitor might have more to do with private events, might be relevant to online human interactions, might cover a wide range of life and might take into account the behaviour of a larger group of people. The concern itself might be irrational but is indeed natural — people care about how they are mirrored even in a purely digital world. The possibility of one's data presence being misused for non-OBA purposes would also add to the problem. The fear of being misunderstood or misrepresented, no matter by insects, animals, humans or machines, is a powerful instinct. In this light, our desire to control our data portrait, even if only faced with machines, would become understandable.

What is even more straightforward is the discriminatory effects of OBA. The fact that the boundary of one's data presence has been blurred can lead to one's employment, healthcare, education and financial opportunities being impaired. Price discrimination is a case in point, and can be generalised to other areas: Without the clear lines between private and public information, between user-given and predicted details, between different databases, and between different users, the targeting of a particular group of people or the sorting of results based on individual differences can be more accurate on discriminatory classifiers, or less accurate on fair classifiers. Either way, discrimination is exacerbated, and less detectable. Also, apart from such potential yet tangible impact on individuals, the long-term filter bubble effect is also problematic. A user's data profile is indeed deduced from their actual activities, which can be arguably quite accurate. However, certain saliences of such presence may be augmented through the processing of an OBA system, which might in turn end up reshaping individual behaviour and trapping

them in the group's own echo chambers. The contours of such bubbles would become even less discernible as a result of the more complicated, interrelated nature of online profiling. People would be less likely to see alternative possibilities in life, let alone to escape such filter bubbles.

3.6.4 Power asymmetries

In Section 3.4, four forms of power imbalance have been flagged up: infrastructural power, informational power, bargaining power and manipulative power. Each of these imbalances would be aggravated when creating a complex advertising network is becoming less costly and more valuable while individual users remain to have few options on how they can negotiate or decide on the use of their data. Tracking, profiling and targeting techniques have had enormous developments since the advent of cookies. At the same time, however, while Internet browsers are becoming more privacy-aware and smarter, the degree to which users are empowered is simply not comparable to the degree to which businesses are empowered. The fluid boundaries of one's data domains allows the online profilers greater flexibility to gain information, to bargain, and to manipulate, whereas individuals are left with greater uncertainties and difficulties to do the same. A lot of average Internet users are even unaware of the fact that they are disempowered during the expansion of the network, not to mention to tackle the problem.

3.6.5 Democracy

Lastly, the democratic issues related to the abuse of online tracking, profiling and targeting can be identified in ways discussed in previous sections. For example, the chilling effect concerning freedom of expression has been explored under the heading of human dignity; the unfair advantages secured by particular political campaigns share the same root with the unfair commercial powers gained by online businesses; the highlighted polarisation

of participants in a civil society may benefit a lot from the discussion of the filter bubble effect. In similar veins, these repercussions will escalate when the vanishing borderlines between data domains become the norm. Such effects are less indirect to individuals but definitely no less important to the society as a whole.

3.7 Summary: From privacy to democracy — Crisis for individuals, collectives and the society

This chapter has so far attempted to cover a wide range of interests that being threatened by big data. They have been examined in the context of OBA, which offers a useful perspective for the clarification and verification of various arguments. In the last section, a theory has been put forward in the search of a coherent account for the risks intensified by big data: It has created a digital reality where there is no longer separation between private and public profiles, between controllable and uncontrollable profiles, between one's profile in one place and another, and between different individuals' profiles. To understand why this metaphor is important, we need to have in mind a picture of how we — as individuals, collectives and a society — have maintained the healthy relationships between each other. In the pre-Internet and the (pre-big data) Internet times, while the walls built on multiple dimensions to maintain obscurity may have drifted from time to time, regulatory or technological solutions have managed to catch up and put them back in place. Today, the walls are not just moving, but rather collapsing.

Such a crisis resulting from the collapse of the walls in the data world should be taken seriously. Throughout this chapter, it has been shown how important it is to maintain these walls for independent individuals, collectives sharing common interests and the civil society. The unclear borderlines of individual data domains against disclosures, against predictions, against

separate data holders, and against others call for further regulatory initiatives to address these new challenges. Before answering the question as to what regulation *can* and *should do*, one would inevitably have to first figure out what regulation *has done*. As such, the focus of the next chapter will be on an important area of regulation over the use of data: data protection law.

Chapter 4 Legal Regulation Confronted with Big Data: Understanding EU Data Protection Law in the Context of Online Behavioural Advertising

4.1 The data protection legal system: An EU perspective

In the previous chapters, a general picture has been sketched out regarding how OBA as a prominent application of big data has connected online users with a wide range of economic actors, as well as how its popularity may bring forth both benefits and risks. Compared to the relatively young trend of big data, data protection law has a history of decades in its course of development. Over the years, the data protection legal framework has been constantly criticised as 'outdated' or 'obsolete', especially in the face of big data.[1] However, it is also believed that, despite the new challenges brought about by big data, the essential objectives and principles of the current data protection regime remain sound.[2] Indeed, it is not unusual that regulatory efforts in certain fields manage to overcome new issues over time, as long as the changes are not fundamental enough, or as long as the safeguards are designed to be future-proof.[3]

---

[1] See, for example, Ira S. Rubinstein, 'Big Data: The End of Privacy or a New Beginning?' (2013) 3(2) International Data Privacy Law 74; Alessandro Mantelero, 'The Future of Consumer Data Protection in the E.U.: Re-thinking the "Notice and Consent" Paradigm in the New Era of Predictive Analytics' (2014) 30 Computer Law & Security Review 643; Steve Lorber, 'Data Protection: A contextual Approach to Regulation' (2014) 14(5) Privacy and Data Protection 11.

[2] 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] OJ L119/1 ('GDPR'), Recital 9.

[3] It is subject to debate whether data protection laws may actually be 'future-proof', but the European Commission seems to believe so. See Commission, 'Proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regards to

Besides, in the realm of data protection, regulatory initiatives have not ceased to move onwards in the hope of keeping pace with the ever-changing socio-technical landscape, at least not in Europe. One of the latest — and most significant — legislative efforts in this area is the General Data Protection Regulation (GDPR), which was adopted on 27 April 2016 and has taken effect as of 25 May 2018.[4] It is well acknowledged by the GDPR that '[r]apid technological developments and globalisation have brought new challenges for the protection of personal data' and thus 'require a strong and more coherent data protection framework in the Union'.[5] Among these developments, big data is probably the most noticeable, and hence, the resolution of its challenges is claimed to lie at the heart of the agenda for this data protection reform.[6] Whether the GDPR can achieve its goal of affording a high level of protection to individuals with regard to their personal data while not stifling data-driven innovation will depend upon the extent to which it can deal with the challenges arising from big data.

To allow for a meaningful assessment of the effectiveness of the GDPR, this chapter will focus on the substantive parts are relevant to OBA — and in a bigger picture, big data — with particular highlights on the newly introduced safeguards. Section 4.2 will briefly present the general data protection framework laid down by the GDPR, in particular how data processing is regulated through the interactions between the basic concepts, essential principles and legitimate grounds. Each of these elements will be analysed in depth in subsequent sections: Section 4.3 addresses the definition of personal

---

the processing of personal data and on the free movement of such data (General Data Protection Regulation)' (2012) COM(2012) 11 final, 104.

[4] GDPR, art 99.

[5] ibid Recital 6.

[6] Commission, *The EU Data Protection Reform and Big Data Factsheet* (2016) <http://ec.europa.eu/justice/data-protection/files/data-protection-big-data_factsheet_web_en.pdf> accessed 22 February 2017.

data and its applicability to the context of OBA; Section 4.4 examines the legal bases by which data processing can be lawfully carried out, with particular emphasis on consent under the GDPR and the ePrivacy Directive; Section 4.5 focuses on three fundamental principles that are particularly relevant to an OBA context; Section 4.6 gives a brief analysis of the implications of the new concept of 'profiling' for OBA. By means of conclusion, Section 4.7 will summarise how the regulatory model has evolved from the DPD to the GDPR, and what legal obstacles the OBA industry will face as the GDPR is enforced.

The reason why this study takes EU data protection law as the legal framework rests on quite a practical consideration. When it comes to data protection law (or known as data privacy law in other jurisdictions), the EU has been a pioneering leader in terms of both legislation and case-law. The availability of abundant materials offers the best opportunity to reflect on the topic with a variety of propositions taken into account. The latest overhaul of data protection law also puts a global spotlight on the latest developments of data protection in Europe. The GDPR, though far from perfect, personifies the up-to-date outcomes of theoretical and practical efforts in promoting data protection over the last two decades, which will be further demonstrated in subsequent sections. Before digging deeper into the specific rules, however, it is essential that we have a general grasp of the history and legal context in which EU data protection law has developed.

4.2 An overview of the General Data Protection Regulation

4.2.1 Regulating data protection in Europe — A very brief history

National initiatives to regulate the use of personal data in Europe date back to the 1970s.[7] Domestic legislation of data protection was first seen in the German

[7] For a more detailed historical overview of national data protection laws in the 1970s, see Judith Rauhofer, 'Privacy Is Dead, Get Over It! Information Privacy and the Dream of a Risk-

federal state Hesse in 1970,[8] followed by nation-wide data protection laws in Sweden (1973),[9] Germany (1977),[10] France (1978),[11] and so on. At European level, the Council of Europe adopted a Resolution in 1973 specifically expressing concerns over the adverse effects of the intensive use of computerised personal information.[12] The Resolution includes a set of principles for protecting individuals from the abuse of data banks in the private sector. Three years later, the work began for drafting the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data.[13] The Convention was open for signature in 1981 and became the first legally binding international instrument with a particular focus on protecting personal data. In addition to expanding and refining the principles already enshrined in the Resolution, certain ground-breaking elements of data protection law were also introduced, including the free movement of personal data across borders and international cooperation.[14] To fully understand the developments of data protection law, it is important to bear in mind that its original aim is to align national laws and eliminate potential obstacles to cross-border data flows.[15] As will be seen below, the objective to achieve a harmonised regulatory regime from which free movement of data may benefit

free Society' (2008) 17(3) Information & Communications Technology Law 185, 188; Eleni Kosta, *Consent in European Data Protection Law* (Brill 2013) 34-73; Gloria González Fuster, *The Emergence of Personal Data Protection as a Fundamental Right of the EU* (Springer 2014) 55-71.

[8] Hessisches Datenschutzgesetz [1970].

[9] Datalag (of Sweden) [1973].

[10] Bundesdatenschutzgesetz (of Germany) [1977].

[11] Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés (of France) [1978].

[12] Council of Europe Committee of Ministers Resolution (73)22 on the Protection of the Privacy of Individuals vis-à-vis Electronic Data Banks in the Private Sector [1973] ('Resolution 73(22)').

[13] Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data [1981] ETS No. 108.

[14] See ibid chs III, IV.

[15] David Bainbridge, *Data Protection* (2nd edn, XPL Law 2005) 16.

has remained a hallmark of subsequent data protection initiatives of international or regional organisations.

Among these initiatives is the OECD Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data in 1980.[16] The Guidelines, though not legally binding on OECD Member countries, have remained influential to date.[17] Eight basic principles were put forward concerning the protection of individual privacy. These principles include collection limitation, data quality, purpose specification, use limitation, security safeguards, openness, individual participation and accountability.[18] The values of utilisation of personal data were also a main concern of the Guidelines, and thus, a different Part also sets out the principles of free flow and legitimate restrictions. National implementation and international cooperation were also touched upon, but only very lightly.

The first international instrument that outlines a full set of regulatory rules and enforcement mechanisms should be accredited to the European Communities, now the EU. Data protection legislation within the Communities was first proposed in 1990 and eventually adopted as Directive 95/46/EC (Data Protection Directive, DPD) in 1995. Inspired both by the Council of Europe Convention and the OECD Guidelines, the DPD incorporates, and adapts, primarily all principles from both instruments. It also remains characterised by the dual objectives of protecting individuals with regard to their personal data, as well as facilitating the free movement of

[16] OECD Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data [1980] ('OECD Guidelines'). The guidelines were updated in 2013, with all eight principles retained in the new version.
[17] Graham Greenleaf, *Asian Data Privacy Laws: Trade and Human Rights Perspectives* (Oxford University Press 2014) 31.
[18] OECD Guidelines, pt 2.

personal data.[19] That latter objective is achieved through the creation of a 'free zone' for data transfers within the EU/EEA where conflict of laws is minimised. [20] Moreover, the concepts of data subject ('an identified of identifiable natural person') and data controller (who 'determines the purposes and means of the processing of personal data'), for instance, are largely the same as defined in the OECD Guidelines.[21] Beyond these inherited features, the DPD includes a larger number of specific definitions and rules to give substance to the fundamental principles.

One of the most significant innovations of the DPD consists in the additional requirement of legitimate grounds for processing of personal data. Under this regime, the processing of personal data in question must comply with *all of* the principles, and also rely on at least *one of* the legal grounds provided by the Directive. In other words, processing of personal data is in principle *prohibited unless* it has a sufficient legal ground that is explicitly provided by law, which is different from some jurisdictions where processing of personal data is generally *allowed unless* otherwise restricted by law. Such legal grounds range from unambiguous consent obtained from the data subject to legitimate interests of the data controller or third parties.[22] That means, data controllers must not just demonstrate that the data protection principles (which are abstract and ambiguous) are fulfilled, but also that they have a sufficient legal basis to carry out the processing.

---

[19] See Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281/31 ('DPD'), Recitals 1-3, art 1. For a detailed analysis, see also Orla Lynskey, *The Foundations of EU Data Protection Law* (Oxford University Press 2015) ch3.

[20] Douwe Korff, *Data Protection Law in Practice in the European Union* (Federation of European Direct and Interactive Marketing and The Direct Marketing Association 2005) 170-171.

[21] See DPD, art 2; OECD Guidelines, pt 1.

[22] DPD, art 7.

On the other hand, the DPD also puts noticeable emphasis on the specific rights of data subjects, such as the right to information,[23] the right of access[24] and the right to object.[25] Procedural and enforcement matters also take up a large part of the DPD, including the registry system,[26] the prior checking requirements,[27] the powers of supervisory authorities[28] and the remedies for data subjects.[29]

The DPD had been such a success in promoting data protection as a fundamental right that, as secondary legislation, it has inspired the incorporation of this right into primary laws of the EU (i.e. the Treaties). In 2000, as the drafting work began for an EU Charter of Fundamental Rights[30], a German representative proposed a separate article for data protection, and it soon gained support from the members of the drafting body.[31] Despite the suggestion by some members that the right to data protection be confined to an aspect of the right to 'respect for private and family life' or even be absorbed into the latter, a separate, stand-alone Article 8 headed 'protection of personal data' eventually made its way into the final text of the Charter.[32] The article has three paragraphs, with emphasis respectively on: a) the official recognition of 'the right to the protection of personal data concerning him or her'; b) the key elements of legitimate use of personal data, namely 'be processed fairly', 'for specified purpose', and safeguarded by 'consent of the person concerned' or 'other legitimate basis', as well as the particular stress on the right of access

[23] ibid ch 2, s IV.
[24] ibid art 12.
[25] ibid art 14.
[26] ibid art 18.
[27] ibid art 20.
[28] ibid art 28.
[29] ibid ch III.
[30] Charter of Fundamental Rights of the European Union [2012] OJ C326/391 ('Charter').
[31] González Fuster (n 7) 195-196.
[32] ibid 196-198.

to, and rectification of, personal data; and c) the oversight of data protection by an independent authority.

It should be noted that the Charter did not have legally binding force until the entry into force of the Treaty of Lisbon in 2009.[33] The Treaty of Lisbon amended two treaties that are now officially known as the Treaty on European Union (TEU) — which gives the Charter 'the same legal value as the Treaties'[34] — and the Treaty on the Functioning of the European Union (TFEU).[35] Article 16 of the TFEU reiterates the right to data protection in almost the same manner as Article 8 of the Charter, and also provides the legal basis for secondary legislation on data protection with no need to articulate the connection with the establishment of the single market.[36] Although the objective of free movement of personal data remains explicitly included in Article 16, greater importance has clearly been given to protecting personal data as a fundamental right at Treaty level.[37] Thus, through Article 8 of the Charter and Article 16 of the TFEU, data protection has in effect gained a 'constitutional' status in the EU legal order, meaning that the interpretation of EU law must be put in line with this right, and any secondary legislation or decision in violation of this right will be found invalid.[38]

As a result of data protection now being an independent fundamental right, the GDPR, unlike the DPD, has shown a complete disconnection from the

---

[33] Treaty of Lisbon amending the Treaty on European Union and the Treaty establishing the European Community [2007] OJ C306/1 ('Treaty of Lisbon').
[34] Consolidated version of the Treaty on European Union [2012] OJ C326/13 ('TEU'), art 6(1).
[35] Consolidated version of the Treaty on the Functioning of the European Union [2012] OJ C326/47 ('TFEU').
[36] González Fuster (n 7) 232-233.
[37] Hielke Hijmans, *The European Union as Guardian of Internet Privacy: The Story of Art 16 TFEU* (Springer 2016) 51.
[38] In fact, the CJEU has already invalidated a number of acts by the EU, notably the Data Retention Directive and the 'Safe Harbor' Decision regarding data transfers to the US. See Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland and Seitlinger and Others* [2014] OJ C 175/6; Case C-362/14 *Schrems* [2015] OJ C 398/5.

terminology of 'privacy', although the regulatory model founded by the DPD has been generally retained. The underlying structure of a set of data protection principles coupled with a list of legitimate bases for data processing remains unchanged. A very brief overview of its principal mechanisms is necessary for the subsequent analysis of their applicability to the practices of OBA.

4.2.2 Concepts, principles, grounds and special rules: The structure of EU data protection law

The GDPR inherits the underlying framework of the DPD in many aspects and remains founded on four sets of building blocks: definitions of key concepts, general data protection principles, legal grounds for processing and special rules in particular contexts. These fundamental ideas are supported by, and materialised as, the rights of data subjects, the responsibilities of data controllers and the enforcement mechanisms implemented by competent authorities. The GDPR itself constitutes a highly sophisticated system that, combined with relevant sector-specific legislation (such as the ePrivacy Directive, as will be analysed in Section 4.4.3), deserves an examination at length. It is however not the task of this study to cover every aspect of the GDPR, nor even some of the most important provisions such as those concerning the public sector.[39] Rather, the focus of this study will be directed towards only those provisions closely related to the case study — the practices of OBA. With this in mind, certain parts of the law would be deliberately omitted or merely lightly touched on. Nevertheless, it is worth a very brief section to go through the four pillars of the GDPR, with general comments on how they relate to the case of OBA. Detailed analyses of the relevant

[39] For example, see DPD, arts 7(e), 8(3) & (4), 26(1)(d).

provisions in the GDPR within the context of the operation of an OBA system will be left to the remaining sections of this chapter.

(a) Definitions

Article 4 of the GDPR defines a series of critical legal concepts and terms, some of which are of paramount significance. For example, the meaning of 'personal data' is defined as 'any information relating to an identified or identifiable natural person ('data subject')'.[40] This concept of personal data is perhaps the most important one — and thus most controversial — because data protection law applies in general only to personal data. If the processing in question is carried out on fully anonymous data, the data protection principles would not apply in the first place.[41] The criterion of identifiability, however, is not always clear-cut.[42] It is often subject to intense debate over whether, say, IP addresses,[43] behavioural[44] or even emotional data[45] of unnamed persons are

[40] GDPR, art 4(1).
[41] ibid Recital 26.
[42] For discussions about the definition of personal data, see Paul M. Schwartz and Daniel J. Solove, 'The PII Problem: Privacy and a New Concept of Personally Identifiable Information' (2011) 86 New York University Law Review 1814; Article 29 Data Protection Working Party, 'Opinion 4/2007 on the concept of personal data' (2007) 01248/07/EN WP 136; Lilian Edwards, 'Taking the "Personal" Out of Personal Data: *Durant v FSA* and its Impact on the Legal Regulation of CCTV' (2004) 1(2) SCRIPT-ed 341; Karen McCullagh, 'Protecting "Privacy" Through Control of "Personal" Data Processing: A Flawed Approach' (2009) 23(1-2) International Review of Law, Computers & Technology 13; Christopher Millard and W. Kuan Hon, 'Defining "Personal Data" in E-social Science' (2012) 15(1) Information, Communication & Society 66.
[43] In the context of ISP use for filtering purposes, the CJEU ruled that IP addresses (and even dynamic IP addresses in some cases) are protected as personal data. See Case C-70/10 *Scarlet Extended* [2012] OJ C 25/6; Case C-582/14 *Breyer* [2016] OJ C 475/3. For further discussions, see Article 29 Data Protection Working Party (n 42) 16-17.
[44] This is also known as 'clickstream data'. For a discussion, see Avi Goldfarb and Catherine E. Tucker, 'Privacy Regulation and Online Advertising' (2011) 57(1) Management Science 57, 60; Thomas M. Lenard and Paul H. Rubin, 'In Defense of Data: Information and the Costs of Privacy' (2010) 2(1) Policy & Internet 149, 164-165.
[45] See Andrew McStay, 'Empathic Media and Advertising: Industry, Policy, Legal and Citizen Perspectives (the Case for Intimacy)' [2016](July-December) Big Data & Society 1; Damian Clifford, 'Citizen-consumers in a Personalised Galaxy: Emotion Influenced

personal data. Some ad network providers have argued that the data they collect and handle is non-personal data, since they cannot tell the identity of the user from the data they process.[46] Privacy advocates, on the other hand, often maintain that such data concerns a person that can be 'singled out' from others, and therefore should be considered personal data and subject to legal restrictions.[47] The GDPR makes particular attempts to mitigate the legal uncertainty arising from the ambiguous scope of personal data. For one thing, Article 4(1) itself, read together with Recitals 26 and 30, provides further guidance on how to determine the personal nature under a given circumstance, including the factors that should be taken into account and a number of instances of these factors.[48] For another, a few new concepts have been introduced to clarify some scenarios that are not clearly covered by the DPD. Pseudonymisation, for instance, is employed to deal with the case where 'the personal data can no longer be attributed to a specific data subject without the use of additional information [..] kept separately […]'.[49] It is made clear that pseudonymised data remains personal data, but the adoption of pseudonymisation can be regarded as a security measure and hence can be used to demonstrate compliance.[50] Another relevant, ground-breaking concept is 'profiling', which refers to the processing of personal data intended

Decision-making, a True Path to the Dark Side?' (2017) 31/2017 CiTiP Working Paper <http://ssrn.com/abstract=3037425> accessed 21 February 2018.

[46] Some marketing associations, for example, claim that '[d]ata about your browsing activity is collected and analysed anonymously'. See EDAA, 'Your Online Choices' <http://www.youronlinechoices.com/uk/about-behavioural-advertising> accessed 26 September 2017; Internet Advertising Bureau UK, *A Guide to Online Behavioural Advertising* (2009) 10 <https://www.iabuk.net/sites/default/files/publication-download/OnlineBehaviouralAdvertisingHandbook_5455.pdf> accessed 7 January 2017.

[47] See Article 29 Data Protection Working Party, 'Opinion 2/2010 on online behavioural advertising' (2010) 00909/10/EN WP 171, 9; Information Commissioner's Office, 'Determining What Is Personal Data' (2012), 8.

[48] GDPR, art 4(1)

[49] ibid art 4(5).

[50] ibid arts 6(4)(e), 25(1), 32(1)(a).

'to evaluate certain personal aspects relating to a natural person'.[51] Such aspects include, *inter alia*, a person's 'personal preferences, interests, […] behaviour, location [and] movements',[52] which are all highly relevant to OBA. These key concepts, as well as a list of others will be further examined as we investigate to what extent the GDPR is applicable to OBA-related activities.

(b) Principles

The GDPR, in its Article 5 ('Principles relating to processing of personal data'), specifies seven principles to which processing of personal data should adhere: (1) lawfulness, fairness and transparency; (2) purpose limitation; (3) data minimisation; (4) accuracy; (5) storage limitation; (6) integrity and confidentiality; and (7) accountability.[53] The essence and implications of the first three principles will be analysed in detail below. One point worth emphasising here, however, is that these principles apply *universally* and *accumulatively*. That means, first, regardless of the purpose of processing (for private or public interest), the justification of processing (consent or otherwise) or the nature of data (sensitive or not, or pseudonymised or not), these principles would generally apply to any processing of personal data so long as it falls within the scope of the GDPR. Therefore, a data controller would nevertheless have to fulfil its obligations under the principle of, for instance, transparency even when the personal data is collected from a third party other than the data subject in a way that does not require consent by the data subject. Second, every single operation conducted on personal data must adhere to *all* of the seven principles, although the GDPR has laid down a limited list of circumstances under which the principles may be restricted.[54] Any non-

[51] ibid art 4(4).
[52] ibid.
[53] ibid art 5.
[54] ibid art 23.

compliance with even just one of these principles would lead to the processing concerned being unlawful.

(c) Legal grounds

Although the data protection principles have been more or less reflected in previous international data protection initiatives (such as the Convention 108 or the OECD Guidelines), it is the EU's data protection law that first lays down the legitimising grounds on which processing of personal data can be based. The DPD provides an enumerative list of six such legal grounds, which are largely retained in the GDPR with certain changes that have greater or lesser significance. These grounds are: (1) the data subject's consent; (2) performance of contract; (3) compliance with legal obligation; (4) vital interests of data subject or another natural person; (5) public interest or official authority; and (6) legitimate interests of data controller or third parties.[55] It should be noted that no basis can be added or removed from this list by Member States.[56] This set of legal grounds functions in a way different from how the general principles do. These legal bases are required in a selective manner, meaning that the data controller only needs to demonstrate the fulfilment of *one* of these requirements. Moreover, a legitimate ground for processing is not required all the time. So long as further processing of personal data is compatible with the initial purposes for which the data is collected, there is no need for that processing to be based on a new legal ground, provided that the initial processing has already been safeguarded by one of the six grounds.[57] Among these grounds, some are obviously more relevant to OBA than others. One might be able to imagine how an ad network provider, for instance, could

[55] ibid art 6.
[56] Joined cases C-468/10 and C-469/10 *ASNEF* [2011] OJ C 25/18 paras 30-32; *Breyer* (n 43) para 57.
[57] GDPR, Recital 50.

claim to justify its practices on the bases of consent, performance of contract and legitimate interests of its own. Yet, it would be very unlikely for them to invoke the grounds of, say, compliance with legal obligation, vital interests of someone or public interests. The latter group of the three bases are more concerned with public policies, not commercial activities. Consequently, the discussion below will be only focused on whether and under what conditions the actors in an OBA network may justify their operations based on the former three grounds.

(d) Special rules

Based on the general framework made up by definitions, principles and legal bases, the specific rules giving substance to the framework constitute the main body of data protection law. Some of these particular provisions have great impact on the operation of OBA businesses. Most notably, Article 21 of the GDPR grants data subjects the right to object to processing for direct marketing purposes.[58] Also, when sensitive data is involved, the processing is prohibited in principle and allowed only in a limited number of exceptional cases such as *explicit* consent having been obtained. Apart from the requirements set out by the GDPR, data controllers would also need to comply with sector-specific *lex specialis*, including, notably for OBA, the ePrivacy Directive. Article 5(3) of that Directive — which will be discussed in further detail in Section 4.4.3 — imposes specific restrictions on the use of tracking techniques like cookies.[59] The consideration of these special rules contributes

[58] ibid art 21.

[59] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) (As amended by Directive 2009/136/EC) [2002] OJ L201/37 ('ePrivacy Directive'), art 5(3).

to a more comprehensive understanding of how OBA is regulated under data protection law.

A last remark of this section is that, the GDPR, or more generally, data protection law, has been developed in an organic manner that cannot be investigated in parts separately. For example, the way 'consent' has been defined has significant impact on what may form valid consent as a ground for data processing. Likewise, the reading of the data minimisation principle would certainly inform the assessment of whether a given case of processing may be allowed for being necessary for the performance of a contract. Therefore, while this chapter will endeavour to analyse the legal framework as clearly as possible, frequent references to other parts of the GDPR, or even other legislation, will be inevitable.

4.3 Defining personal data through identifiability: Is online behavioural data personal data?

It appears to be a straightforward question as to whether the data collected by OBA platforms in the tracking phase should be considered personal data. If the data is all *about* what the users have done on the Internet, how can it be possible that such data is not *their* personal data? This should not be a difficult problem, in particular not so under the GDPR, as we will see later. However, at the time of the DPD, there was indeed a legal debate between the industry and the regulators.[60] The point of dispute lies in the concept of identifiability: Are Internet users identifiable when they are being tracked by an OBA system?

The problem results partly from the insufficient clarity of the 'personal data' definition under the DPD. Its Article 2(a) conceptualises personal data as 'any

[60] For example, see Ronald Leenes, 'Do They Know Me? Deconstructing Identifiability' (2007) 4(1&2) University of Ottawa Law & Technology Journal 135; Eric Picard, 'The Ethical Issues with 3rd Party Behavioral Tracking' *AdExchanger* (31 October 2011) <https://adexchanger.com/the-debate/3rd-party-behavioral-tracking/> accessed 26 September 2017.

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 identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity'.[61] More clarifications have been given in Recital 26, which states that, 'to determine whether a person is identifiable, account should be taken of all the means likely reasonably to be used either by the controller or by any other person to identify the said person'.[62]

According to the logic of the DPD, for a set of data to be personal data, it needs to be associated with the identity or certain characteristics of a person. This can be done with the help of additional information that can be reasonably acquired by the data controller or a third person. The 'additional information' requirement turns out to hold particular significance at a time of big data, when large volumes of data that was thought 'anonymised' can in fact be 'de-anonymised' by bringing together pieces of information from various sources. It is well documented how this can be done with striking ease,[63] and in this light, a much wider scope of data would fall within the definition of personal data as long as it has the potential to contribute to the 'database-in-the-sky' that can connect all the dots about anyone.[64]

Even setting aside the complexities of integrated data, the identifiable nature of a piece of information alone can be intractable. The identity (or the personal features) of an identifiable person would be of a kind either explicitly

[61] DPD, art 2(a).
[62] ibid Recital 26.
[63] Plenty of examples of how reidentification of anonymised data can be and has been achieved can be found in Paul Ohm, 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' (2010) 57 UCLA Law Review 1701; Schwartz and Solove (n 42); Manon Oostveen, 'Identifiability and the Applicability of Data Protection to Big Data' (2016) 6(4) International Data Privacy Law 299.
[64] Ohm (n 63) 1748.

designated by law (e.g. 'identification number') or one that is distinctive enough to specify the person (e.g. 'physical, physiological, mental, economic, cultural or social identity'). The distinction between these two ways of establishing 'identifiability' is noted by the Article 29 Working Party, who distinguish 'directly identifiable' from 'indirectly identifiable' persons.[65] According to their analysis, a person's name, as a most common identifier, would be precise enough to ascertain the identity of a person in many contexts.[66] A person's identification number, by the same token, would generally suffice as a direct identifier. As regards indirect identifiers, if various categories of information about a person's certain characteristics come together to form a 'unique combination', that person would be considered identifiable.[67]

In the case of OBA, however, it is not entirely clear whether the data collected by ad network providers should *always* be considered personal data under the DPD. One who wishes to argue that such data constitutes personal data could have two approaches: First, to argue that the behavioural or demographic data contained in the datasets is specific enough to identify the individual; second, to argue that the dataset contains a particular form of 'identification number'. However, the first approach appears to have raised more questions than it has solved. Again, what would be precise enough to tell a person from others is not free from disputes.[68] For a non-registered user, the functioning of an OBA system would probably not involve the real name, ID number or address of a user, but the system is very likely to have some

[65] Article 29 Data Protection Working Party, 'Opinion 4/2007 on the concept of personal data' (n 42) 12-15.
[66] ibid 13. Interestingly, however, the only type of identifier explicitly listed by Article 2(a) of the DPD is 'identification number', not 'name'.
[67] ibid.
[68] Edwards (n 42); McCullagh (n 42).

clues about the user's neighbourhood, gender or consumption preferences.[69] Would these factors count as 'physical, physiological, mental, economic, cultural or social identity'? The law itself has not provided further answers.

An alternative approach is to prove that the tracking techniques employed in OBA involve an 'identification number' or equivalent, so there is no need to further show that the person can be identified from other features. Of course, rarely would any ad network provider collect the national ID number of their users, but what about the unique ID that they assign to each user and store/retrieve as cookies, or other identifiers like IP addresses? If these forms of device fingerprinting [70] can be considered 'identification number' or something to that effect, then there would be no doubt that such data is personal data. The Article 29 Working Party seems to have taken this strategy. It is believed that 'behavioural advertising normally involves the collection of IP addresses and the processing of unique identifiers (through the cookie). The use of such devices with a unique identifier allows the tracking of users of a specific computer even when dynamic IP addresses are used.'[71]

For this reason, during the legislative process of the GDPR, the Working Party insisted that certain recitals of the proposal be redrafted to explicitly include the situations where the data subject can be 'singled out' by means of cookies, IP addresses or otherwise.[72] As a result, in the final text of the GDPR, it is stated in Recital 30 that '[n]atural persons may be associated with online

[69] See, for example, Google, 'Data Collection | How Google Uses Your Personal Information' <https://privacy.google.com/intl/en-GB/your-data.html> accessed 7 March 2017.
[70] For the concept of device fingerprinting, see Article 29 Data Protection Working Party, 'Opinion 9/2014 on the application of Directive 2002/58/EC to device fingerprinting' (2014) 14/EN WP 224.
[71] Article 29 Data Protection Working Party, 'Opinion 2/2010 on online behavioural advertising' (n 47) 9.
[72] Article 29 Data Protection Working Party, 'Opinion 01/2012 on the data protection reform proposals' (2012) 00530/12/EN WP 191, 9-10.

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.'[73] Accordingly, the explanation in Article 4(1) regarding the meaning of 'an identifiable natural person' has been amended as 'one who can be identified, directly or indirectly, in particular by reference to an identifier such as a *name*, an identification number, *location data*, an *online identifier* or to one or more factors […]'.[74] The GDPR has thereby expanded the list of 'direct identifiers' from only 'identification number' to three more categories, including online identifiers. Read together with Recital 30, which clearly states that cookie identifiers are one form of online identifier, there would be little, if any, room for the argument that data collected through tracking cookies, or any other forms of identifiers created by individual- or device-specific tracking techniques, does not constitute personal data. This clearly marks a paradigm shift from the DPD whose concerns are almost entirely about the real-life identities, to the GDPR which underlines identifiability also in terms of online identities. As illustrated in Chapter 1, tracking (by cookies or similar techniques) serves as the foundation stone of an OBA system, and any behavioural information is worthless unless associated with a particular tracking id. To such an extent, it should be concluded that, under the GDPR, processing of online behavioural data in OBA systems is fully subject to data protection law.

[73] GDPR, Recital 30.
[74] ibid art 4(1) (emphasis added).

4.4 Legal grounds: What legal basis for OBA use of personal data?

4.4.1 Consent, contract and legitimate interests

As we have seen above, the GDPR (like the DPD) provides six legal grounds on which the data controller may legitimise the processing of personal data: (1) consent; (2) performance of contract; (3) compliance with legal obligation; (4) vital interests of data subject or another natural person; (5) public interest or official authority; and (6) legitimate interests of data controller or third parties.[75] Depending on the specific circumstances, any of these legal bases might be relevant for an instance of data processing by an OBA operator. For the private sector, however, three of these bases are much more relevant for their day-to-day business. For instance, a retailer may use their customers' address details for the delivery of goods under a sales contract.[76] With regard to OBA, it is not fully clear which one(s) could constitute a valid legal basis for the use of personal data.

Zuiderveen Borgesius has carried out a thorough analysis of the applicability of these three bases to OBA, and concluded that the only practical basis is consent.[77] Although the analysis is conducted within the legal framework of the DPD, the conclusion about the grounds of 'performance of contract' and 'legitimate interests' remains sound as no material change has been made by the GDPR regarding these two grounds.

---

[75] ibid art 6(1).

[76] 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' (2014) 844/14/EN WP 217, 17; Article 29 Data Protection Working Party, 'Guidelines on consent under Regulation 2016/679' (2018) 17/EN WP259 rev.01, 9.

[77] Frederik J. Zuiderveen Borgesius, 'Personal Data Processing for Behavioural Targeting: Which Legal Basis?' (2015) 5(3) International Data Privacy Law 163. See also Frederik J. Zuiderveen Borgesius, *Improving Privacy Protection in the Area of Behavioural Targeting* (Kluwer Law International 2015) 147-186.

The reason why an OBA operator cannot rely on the ground of necessity for performance of contract is that, even in the case where the user has indeed entered into a contract with the service provider (which is not always the case), it would be almost impossible for the data controller to prove such processing for OBA purposes to be 'necessary' for the performance of the contract.[78] Similarly, the Working Party has taken the view that the performance of contract is not an appropriate basis for profiling based on online behavioural data, whether such processing is covered by a contract or not.[79] What matters, as the Working Party argues, is 'the exact *rationale* of the contract, i.e. its substance and fundamental objective'.[80]

By the same token, neither the 'legitimate interests' of the data controller nor third parties would be a viable option as the industry can barely pass the necessity threshold.[81] It is argued that, while direct marketing is indeed a legitimate interest,[82] alternative business and technical models that are less intrusive, such as contextual advertising or browser-based profiling, may render the current OBA practices disproportionate.[83]

For these reasons, the safest, if not the only, legal basis for OBA operators to count on is consent. This is also the view taken by the Working Party in a 2013 Opinion.[84] Another reason — perhaps the most practical one — for this

---

[78] Zuiderveen Borgesius, 'Personal Data Processing for Behavioural Targeting: Which Legal Basis?' (n 77) 165-167.
[79] 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 76) 17.
[80] ibid (emphasis original).
[81] Zuiderveen Borgesius, 'Personal Data Processing for Behavioural Targeting: Which Legal Basis?' (n 77) 168.
[82] As a matter of fact, the GDPR has expressly recognised direct marketing as a legitimate interest. See GDPR, Recital 47.
[83] Zuiderveen Borgesius, 'Personal Data Processing for Behavioural Targeting: Which Legal Basis?' (n 77) 168.
[84] Article 29 Data Protection Working Party, 'Opinion 03/2013 on purpose limitation' (2013) 00569/13/EN WP 203, 46.

study to focus only on consent is that, as will be seen later in this chapter, consent is the only legal ground for processing of personal data collected by cookies.[85] In other words, it is safe to conclude that consent is the only legitimate basis for data processing in the tracking stage. As regards the profiling and targeting stages, while in theory performance of contract and legitimate interest may arguably be acceptable as the valid legal basis, the analysis above shows that they are not practical options. The legal requirements surrounding consent will be further discussed in subsequent sections, in the light of the GDPR and the ePrivacy framework.

4.4.2 Consent in the GDPR

Consent is the first — and sometimes considered the most important — basis provided by data protection law.[86] The prevalence of this legal basis perhaps results from its nature of being most indigenous to the idea of respecting the data subject's choice. This is also the most controversial one, partly as a result of an arguable over-reliance on this basis by data controllers in the past,[87] and partly as a result of the uncertainty of the validity of consent in practice.[88] Article 4(11) lays down a number of restrictions by defining consent as 'any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative

[85] ePrivacy Directive, art 5(3).
[86] Article 29 Data Protection Working Party, 'Opinion 15/2011 on the definition of consent' (2011) 01197/11/EN WP187, 7. See also TFEU, art 16(2).
[87] Bart W. Schermer, Bart Custers and Simone van de Hof, 'The Crisis of Consent: How Stronger Legal Protection May Lead to Weaker Consent in Data Protection' (2014) 16 Ethics and Information Technology 171; Judith Rauhofer, 'Of Men and Mice: Should the EU Data Protection Authorities' Reaction to Google's New Privacy Policy Raise Concern for the Future of the Purpose Limitation Principle?' (2015) 1(1) European Data Protection Law Review 5, 14.
[88] This is in particular the case when the use of cookie is involved. See Andrew McStay, 'I Consent: An Analysis of the Cookie Directive and Its Implications for UK Behavioral Advertising' (2012) 15(4) New Media & Society 596.

action, signifies agreement to the processing of personal data relating to him or her'.[89] Several elements can be easily identified from the text of this provision.

(a) 'freely given'

First of all, consent must be 'freely given'. While the DPD does not provide the criteria as to what does or does not constitute 'freely given' consent, the GDPR clarifies that '[c]onsent should not be regarded as freely given if the data subject has no genuine or free choice or is unable to refuse or withdraw consent without detriment.'[90] This echoes the Working Party's long-standing position that 'if the data subject has no real choice, feels compelled to consent or will endure negative consequences if they do not consent, then consent will not be valid.'[91] It is therefore clear that if the data subject's refusal to give consent would put them into an adverse situation, then the consent in dispute could not be considered freely given. How 'negative' should be gauged, however, is not without doubts. A public authority threatening to issue a fine or an employer firmly demanding the disclosure of personal data are of course examples of consent *not* being obtained free from duress, because the data subject would be worse off than their *status quo* should they refuse to consent.[92] Yet, the more difficult question is, what if the consequence of not giving consent is simply the loss of interests or opportunities from which the data subject could have benefited? What if, for example, an Internet user would be denied service by a website unless they agree to allow the website to collect

[89] GDPR, art 4(11).
[90] ibid Recital 42.
[91] Article 29 Data Protection Working Party, 'Guidelines on consent under Regulation 2016/679' (n 76) 5. See also Article 29 Data Protection Working Party, 'Opinion 15/2011 on the definition of consent' (n 86) 12.
[92] Article 29 Data Protection Working Party, 'Guidelines on consent under Regulation 2016/679' (n 76) 6-7.

their personal data in the first place? Perhaps this can be examined with regard to the 'without detriment' requirement highlighted in Recital 42. The Working Party takes the view that, if the withdrawal of consent would lead to 'the performance of the service being downgraded to the detriment of the user', then 'consent was never validly obtained'.[93]

Article 7(4) sheds further light on this problem. It reads, '[w]hen assessing whether consent is freely given, utmost account shall be taken of whether, *inter alia*, the performance of a contract, including the provision of a service, is conditional on consent to the processing of personal data that is not necessary for the performance of that contract.'[94] In other words, if the processing of personal data is not necessary for service provision yet the provider nevertheless requests it as a precondition, then the consent would be very likely to be found invalid, as that situation would be taken into 'utmost account'. When the draft GDPR was considered by the European Parliament, an even tougher version was proposed, requiring that '[t]he execution of a contract or the provision of a service shall not be made conditional on the consent to the processing of data that is not necessary for the execution of the contract or the provision of the service […].'[95] This was first entirely rejected by the Council,[96] but then later reintroduced with a softer tone[97] and rephrased

[93] ibid 11.
[94] GDPR, art 7(4).
[95] European Parliament, 'Report on the proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)' (2013) A7-0402/2013, 72.
[96] Council, 'Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) - Preparation of a general approach' (2015) ST 9565 2015 INIT, 85.
[97] Council, 'Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) - Chapter II, preparation of trilogue' (2015) ST 11245 15 INIT, 7.

to its current form as the European Parliament insisted on the strengthened wording.[98] The current toned-down version means that prerequisite consent to data processing unnecessary for the performance of contract is not *always* invalid, but the data controller would need to demonstrate a very strong justification for this.[99] As such, for the acquired consent to be valid, OBA operators must prove either that those users who refuse to consent would not be denied service, or that there is a compelling reason to outweigh the 'utmost account' that would otherwise lead to the conclusion that the consent is not 'freely given'.

In essence, the effect of Article 7(4) is to unbundle consent to data processing from the underlying service contract, which leads to the practical question regarding the substantial difference between these two consensual bases. In theory, of course, consent is no different from a contract to the extent that it represents the mutual agreement between the data subject and data controller on how personal data should be processed. Indeed, the Working Party suggests that when determining the validity of consent, the requirements laid down by other areas of law, such as contract law, should be taken into account.[100] Yet, in cases where such consent has not been obtained at the time of the conclusion of the contract, the data controller may still

---

[98] Council, 'Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) - Written debriefing of trilogue on 24 November' (2015) ST 14461 15 INIT, 2.

[99] It is still open to debates what constitutes a justification strong enough, but it is generally believed that the conditions should be strictly interpreted. See Information Commissioner's Office, 'Consultation: GDPR Consent Guidance' (2017), 19-21; Philipp Hacker, 'Personal Data, Exploitative Contracts, and Algorithmic Fairness: Autonomous Vehicles Meet the Internet of Things ' (2017) 7(4) International Data Privacy Law 1, 10-11; Bojana Kostić and Emmanuel Vargas Penagos, 'The Freely Given Consent and the "Bundling" Provision under the GDPR' (2017) 153 Computerrecht 217.

[100] Article 29 Data Protection Working Party, 'Opinion 15/2011 on the definition of consent' (n 86) 6.

process the data provided that this proves to be necessary for the fulfilment of contractual or pre-contractual arrangements. The real problem, however, is whether data controllers may use the contracts with their registered users — e.g. terms and conditions — to cover the processing of personal data for OBA purposes.

Article 7(4) has clearly envisaged the possibility that consent co-exists with a contract while the former does not necessarily form part of the latter. Indeed, the Article is intended to regulate the bundling of these two measures by questioning the validity of such a practice. In this light, it is sometimes even necessary to hold them apart while they are formally intermingled, so as to avoid circumvention of this requirement. This is in line with the Working Party's position that 'Article 7(4) seeks to ensure that the purpose of personal data processing is not disguised nor bundled with the provision of a contract of a service for which these personal data are not necessary.'[101] This is further illustrated with a use case of mobile app: Processing of location data for OBA purposes is not necessary for the functioning of a photo-editing app and thus access to the app cannot be made conditional on consent to the use of such data.[102] That means, the demarcation of the scope of a 'contract' involves substantive, not formal, assessment. If the processing concerned is found unnecessary for the genuine, principal purpose of a contract (e.g. provision of social media service), the authorisation given by the data subject regarding the processing in question, whether formulated as part of a written contract or not, would be considered consent, not a contract. This way, the validity of such

[101] Article 29 Data Protection Working Party, 'Guidelines on consent under Regulation 2016/679' (n 76) 9. See also Article 29 Data Protection Working Party, 'Opinion 15/2011 on the definition of consent' (n 86) 8; Frederik J Zuiderveen Borgesius and others, 'Tracking Walls, Take-It-Or-Leave-It Choices, the GDPR, and the ePrivacy Regulation' (2017) 3(3) European Data Protection Law Review 353, 360-361.

[102] Article 29 Data Protection Working Party, 'Guidelines on consent under Regulation 2016/679' (n 76) 6.

authorisation is subject to rules governing consent, such as Article 7(4), which considers such consent generally not 'freely given' if the service is provided on condition of giving consent. The Working Party takes the view that OBA is a different purpose separable from — and unnecessary for — the provision of online services.[103] In other words, the practice of 'tracking wall' — that is, requiring users to consent to the use of their data for advertising purpose before they can access the service — will be likely to be considered incompatible with the GDPR, at least from the regulators' perspective.[104]

(b) 'specific'

A valid instance of consent also needs to be 'specific'. Unlike the 'freely given' element, the GDPR itself has not provided any clarification or examples on what amounts to specific consent. Therefore, further guidance has to be sought from opinions issued by regulators. In its Opinion on the definition of consent under the DPD, the Working Party provides its interpretation that '[t]o be specific, consent must be intelligible: it should refer clearly and precisely to the scope and the consequences of the data processing.'[105] It is further explained that 'open-ended' consent cannot be specific, and the context in which the consent applies must be limited.[106] Such a context should be specified notably on these aspects: the scope of data being processed, the purposes of the processing, and the activities or consequences of such processing.[107] These requirements are also reflected in Articles 13 and 14, which lay down the

[103] 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 76) 17; Article 29 Data Protection Working Party, 'Guidelines on consent under Regulation 2016/679' (n 76) 7-10.
[104] Zuiderveen Borgesius and others (n 101) 360-361.
[105] Article 29 Data Protection Working Party, 'Opinion 15/2011 on the definition of consent' (n 86) 17.
[106] ibid.
[107] ibid.

categories of information that should be provided to the data subject, although the scope of information covered by these two Articles is wider than what constitutes 'specific' consent. For this reason, the 'specific' requirement is intrinsically connected to the 'informed' element as will be discussed below.[108]

Apart from ensuring the context of consent is not too general, the Working Party has also added one more dimension to the specificity criteria in its latest guidelines regarding the GDPR.[109] Granularity, the Working Party argues, underlies not just the 'freely given' test but also the 'specific' one. In other words, data processing for multiple purposes may not be covered by one all-inclusive instance of consent, but should each be justified by specific consent.[110] The unbundling effect here is thus slightly different from that stemming from the 'freely given' condition: While the 'freely given' element deals mainly with the conditionality imposed between the main service and additional data processing, the 'specific' element can further question the specificity of such additional processing. Taking OBA as an example, it is a matter of voluntariness whether a social media service provider may use one instance of consent to cover both purposes of 'provision of service' and 'direct marketing'. Yet, it is a different matter — a matter of specificity — whether the service provider may use one instance of consent to cover, say, both 'commercial marketing' and 'political marketing'. The desired level of granularity under this new approach to the 'specific' requirement — which, as will be analysed below, also concerns the principle of purpose limitation — remains highly unclear at this point. Under the DPD, Korff suggests that in general cases a statement of 'for direct marketing' would be specific enough.[111]

---

[108] ibid.
[109] Article 29 Data Protection Working Party, 'Guidelines on consent under Regulation 2016/679' (n 76) 12.
[110] ibid 13.
[111] Korff (n 20) 35.

In various documents and recommendations, the Working Party also seems to show approval of the wording of this purpose when they consistently use the expression of 'for direct marketing purposes'.[112] Despite the same (and equally ambiguous) wording under both the DPD and the GDPR, it is at least disputable — in particular considering the state-of-the-art techniques developed by the marketing sector in recent years — whether the mere reference to 'direct marketing' would be specific enough any longer under the GDPR.

(c) 'informed'

The third criterion of the validity of consent is whether it is given in an 'informed' manner. The Working Party explains that, '[f]or consent to be informed, it is necessary to inform the data subject of certain elements that are crucial to make a choice' and such information 'must be clear and distinguishable from other matters and provided in an intelligible and easily accessible form.'[113] As such, the 'informed' requirement is directly linked to the principles of transparency and purpose limitation, as well as the data controller's obligation to provide information. A close connection and even overlap thus clearly exists between the 'specific' and 'informed' requirements.[114] Indeed, Recital 42 of the GDPR makes it clear that '[f]or consent to be informed, the data subject should be aware at least of the identity

[112] For example, see Article 29 Data Protection Working Party, 'Opinion 15/2011 on the definition of consent' (n 86); Article 29 Data Protection Working Party, 'Opinion 03/2013 on purpose limitation' (n 84); 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 76).

[113] Article 29 Data Protection Working Party, 'Guidelines on consent under Regulation 2016/679' (n 76) 14.

[114] In the public consultation on its guidelines on consent under the GDPR, the UK ICO proposes an analysis that puts both requirements under one section. See Information Commissioner's Office, 'Consultation: GDPR Consent Guidance' (n 99) 21-22.

of the controller and the purposes of the processing for which the personal data are intended',[115] which echoes the discussion of specificity above.

Having said that, as pointed out before, the information that a data controller is obliged to provide has a wider scope than what the 'specific' component requires. In accordance with Articles 13 and 14 of the GDPR, the data controller should provide, among other things, the following details: (a) the information about the data controller; (b) the purposes of, and the legal basis for, the processing; (c) the categories of personal data concerned; (d) the identities or categories of the data recipients; (e) the period of storage of personal data; (f) the data subject's right to withdraw consent; and (g) the existence of automated decision-making, the logic involved and the potential impacts upon the data subject.[116]

As far as an OBA system is concerned, all these aspects should be covered by the privacy policy or consent statement when the consent is obtained. Apart from specifying the purposes, which has been analysed in the preceding section, the information about the data controllers is also indispensable. As shown in Chapter 1, the collection of personal data within an OBA network usually takes place not just in the website that the user intends to visit (the 'ad publishers') but also the servers of ad network providers and a few more categories of actors. Such information as to the scope of the entities who are involved in the data collection and subsequent dissemination (i.e. the advertisers, ad publishers, ad network providers and other involved parties) should be clearly stated to the data subject. Equally important is the scope of data being processed (i.e. cross-site behavioural and demographic data) as

[115] GDPR, Recital 42.
[116] ibid arts 13, 14.

well as the logic of how data is processed and what kind of potential consequences might result from such processing.

(d) 'unambiguous'

The last defining element of the notion of consent — 'unambiguous' — might sound like a synonym for the 'specific' requirement but in fact has different implications. The condition of unambiguity requires that the action by the data subject must represent a clear indication of their will. If it cannot be ascertained that the data subject indeed means to authorise the specific processing concerned, the consent cannot be considered 'unambiguous'. For this reason, the GDPR itself spells out quite clearly that, by way of example, '[t]his could include ticking a box […], choosing technical settings […]', whereas '[s]ilence, pre-ticked boxes or inactivity' would not meet the criterion.[117] According to this logic, one of the most popular practices on the Internet that treats users' navigation through the website as giving consent may very well fail to pass the unambiguity test: Arguably, the mere clicking on a link on a webpage with a warning banner conveys an even weaker message of acceptance than by clicking an 'I agree' button along with a pre-ticked box. This also mirrors the consistent position of the Working Party (even before the adoption of the GDPR) that, with the example of an online game, the indication of intention by using the website and providing personal details does not constitute unambiguous consent for processing player data for marketing purpose.[118] In the latest guidelines on consent, it is made even clearer that 'merely continuing the ordinary use of a website is not conduct from which one can infer an

[117] ibid Recital 32.

[118] Article 29 Data Protection Working Party, 'Opinion 15/2011 on the definition of consent' (n 86) 23.

indication of wishes by the data subject to signify his or her agreement to a proposed processing operation.'[119]

(e) 'explicit' (?)

The wording of 'explicit' is not explicitly included in the definition of consent, but the phrase 'explicit consent' appears in various provisions such as consent for processing of sensitive data,[120] and for automated decision-making.[121] Prior to the adoption of the GDPR, the distinction between 'unambiguous consent' (for non-sensitive data)[122] and 'explicit consent' (for sensitive data)[123] had already caused a degree of confusion.[124] In the original Commission proposal of the GDPR, it was suggested that '[i]n the definition of consent, the criterion "explicit" [be] added to avoid confusing parallelism with "unambiguous" consent and in order to have one single and consistent definition of consent, ensuring the awareness of the data subject that, and to what, he or she gives consent.'[125] Accordingly, consent was defined in that draft as 'any freely given specific, informed and *explicit* indication'.[126] The same approach gained support from the European Parliament in the first reading.[127] However, this proposed definition suffered a significant setback when being considered by the Council: 'At the COREPER meeting of 8 May 2013, many delegations stated that the requirement for "explicit" consent in all cases — which differ

[119] Article 29 Data Protection Working Party, 'Guidelines on consent under Regulation 2016/679' (n 76) 17.
[120] GDPR, art 9(2)(a).
[121] ibid art 22(2)(c).
[122] ibid art 7(a).
[123] ibid art 8(1).
[124] It is even suggested that the 'unambiguous' qualification is redundant as it adds no value to the interpretation of valid consent. See Kosta (n 7) 235.
[125] Commission, 'Proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regards to the processing of personal data and on the free movement of such data (General Data Protection Regulation)' (n 3) 8.
[126] ibid 42 (emphasis added).
[127] European Parliament (n 95) 65.

[*sic*] from the requirements of the 1995 Data Protection Directive — was unrealistic.'[128] The 'explicit' requirement was then removed from the draft[129] — on which only the Commission and seven Member States made a reservation [130] — and was, again, substituted with the 'unambiguous' phraseology (as in the DPD) in almost the last minute of the trilogue, to ensure what was called a 'balance'.[131]

Hence, the legislators have clearly made a distinction between '(unambiguous) consent' and 'explicit consent', and required that sensitive data be treated with special protection. This differentiation, however, does not mean that processing of non-sensitive data can be justified by what is sometimes referred to as 'implied consent' or 'opt-out consent' by the industry[132] and even regulators[133]. As explained above, the 'unambiguous' requirement itself has precluded the possibility of pre-ticked boxes, inactivity or similar approaches as valid consent. This interpretation is also suggested in

[128] Council, 'Proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) - Agreement on key issues of Chapters I-IV' (2013) ST 9398 13 INIT, 6.
[129] ibid 49.
[130] Council, 'Proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)' (2013) ST 11013 13 INIT, 62.
[131] Council, 'Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) [first reading] - Analysis of the final compromise text with a view to agreement' (2015) ST 15039 2015 INIT, 3.
[132] For example, see Internet Advertising Bureau UK, *IAB Affiliate Marketing Council Consumer Transparency Framework: A Guide for Publishers / Affiliates* (2013) 5 <https://iabuk.net/sites/default/files/Consumer%20Transparency%20Framework%20V1.2.pdf> accessed 7 March 2017.
[133] For example, see Information Commissioner's Office, 'Guidance on the rules of cookies and similar technologies' (2012) V. 3, 6. Please however note that this guidance is intended to address issues arising from the use of cookie, which, as pointed out above, is subject to a different set of rules governing the general use of personal data. However, the ICO has not suggested that the possibility of 'implied consent' is limited only to the scenario involving cookies.

a latest document of the Working Party, in which the 'explicit' requirement is considered an extra safeguard (such as two-stage verification) in addition to the already escalated 'unambiguous' standard.[134] The legitimacy of the so-called 'implied' or 'opt-out' consent will be further analysed in the next section regarding consent for 'cookies', since such practices are often contended for in that context.

4.4.3 Consent in the ePrivacy Directive

(a) The development and enforcement of the so-called 'cookie law'

The GDPR — or its predecessor, the DPD — is not the only data protection statute at EU level. In fact, certain sector-specific legislations exist, with most notably the ePrivacy Directive dealing with data protection matters in the electronic communications sector.[135] Although the operation of OBA systems usually does not involve the provision of publicly available electronic communications services in public communication networks,[136] one provision of the ePrivacy Directive holds special significance to OBA practices. Techniques that store or retrieve information from a networked terminal equipment, particularly cookies and other forms of device fingerprinting, are regulated by Article 5(3) of the ePrivacy Directive.

In the original 2000 Commission proposal of the Directive, Article 5 ('Confidentiality of the communications', which remains unchanged in later versions) mainly tackles only 'listening, tapping, storage or other kinds of

[134] Article 29 Data Protection Working Party, 'Guidelines on consent under Regulation 2016/679' (n 76) 18.
[135] ePrivacy Directive.
[136] By virtue of deep-packet inspection, ISPs can actually assist advertising, but this is not considered in this study. For further details, see Paul Ohm, 'The Rise and Fall of Invasive ISP Surveillance' [2009] University of Illinois Law Review 1417.

interception or surveillance'.[137] During the first reading by the European Parliament, a new paragraph was proposed: 'Member States shall prohibit the use of electronic communications networks to store information or to gain access to information stored in the terminal equipment of a subscriber or user without the prior, explicit consent of the subscriber or user concerned.'[138] This new provision was justified by an explanation that:

> '[t]erminal equipment of users of electronic communications networks and any information stored on such equipment are part of the private sphere of the users […]. So-called cookies, spyware, web bugs, hidden identifiers and other similar devices that enter the users' terminal equipment without their explicit knowledge or explicit consent […] may seriously intrude the privacy of these users.'[139]

The stringent condition of 'prior, explicit consent', however, was knocked down under the consideration of the Council in the face of industrial opposition,[140] and was replaced with a much softer version that allowed cookies 'on condition that the subscriber or user concerned receives in advance clear and comprehensive information […] and is offered the right to refuse such processing by the data controller.'[141] This approach was backed by the Commission, who considered it a 'right balance between the EP amendment and the concerns that were raised by economic operators regarding that

---

[137] Commission, 'Proposal for a Directive of the European Parliament and of the Council concerning the processing of personal data and the protection of privacy in the electronic communications sector' (2000) COM/2000/0385 final.

[138] European Parliament, 'Second Report on the proposal for a European Parliament and Council directive concerning the processing of personal data and the protection of privacy in the electronic communications sector ' (2001) A5-0374/2001, 22.

[139] ibid.

[140] Eleni Kosta, 'Peeking into the Cookie Jar: The European Approach Towards the Regulation of Cookies' (2013) 21(4) International Journal of Law and Information Technology 380, 387.

[141] Council, 'Common Position adopted by the Council on 28 January 2002 with a view to the adoption of a Directive of the European Parliament and of the Council concerning the processing of personal data and the protection of privacy in the electronic communications sector' (2002) 15396/2/01 REV 2.

amendment.' [142] Interestingly enough, when this common position was returned to the Parliament for a second reading, the wording had been even further toned down by the Parliament from 'receives in advance [information]' to 'has access to'.[143] The Parliament adopted a dramatic change of position: 'Cookies are legitimate tools which serve a range of useful purposes […] In addition, the means to accept and/or reject cookies already exist in most browser software. Consequently, the obligation for website operators to provide this possibility is superfluous.' [144] With some slight changes of language, the 'clear and comprehensive information' + 'right to refuse' model — sometimes known as 'opt-out consent' — was eventually adopted in the ePrivacy Directive in 2002.

Five years later, in 2007, the Commission proposed a legal reform on multiple legislations, including the ePrivacy Directive, with a view to improving protection to consumers and users in an information society.[145] While Article 5(3) was part of the package, the proposed amendment was largely a definitional patch on the scope of that provision,[146] with the

---

[142] Commission, 'Communication from the Commission to the European Parliament pursuant to the second subparagraph of Article 251 (2) of the EC Treaty concerning the common position of the Council on the adoption of a Directive of the European Parliament and of the Council on processing of personal data and the protection of privacy in the electronic communications sector' (2002) SEC/2002/0124.

[143] European Parliament, 'Recommendation for Second Reading on the Council common position for adopting a European Parliament and Council directive concerning the processing of personal data and the protection of privacy in the electronic communications sector' (2002) A5-0130/2002, 13.

[144] ibid.

[145] Commission, 'Proposal for a Directive of the European Parliament and of the Council amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on consumer protection cooperation' (2007) COM(2007) 698 final.

[146] For more background information, see Kosta, 'Peeking into the Cookie Jar' (n 140) 383-385.

regulatory model entirely unchanged.[147] The idea of switching the condition for using cookies from 'opt-out consent' to 'opt-in/prior consent' was first raised by the Parliament during first reading.[148] It should however be noted that the Parliament had also amended a Recital to recognise that 'browser settings constitute prior consent',[149] which received strong objection from the Working Party.[150] Both the 'prior consent' requirement and the recital concerning browser settings were squarely rejected by the Commission in its amended proposal[151] as well as by the Council in its common position.[152] In the course of second reading, such amendments were brought up again by the Parliament, although with a much less forthright tenor on the point of browser

[147] Commission, 'Proposal for a Directive of the European Parliament and of the Council amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on consumer protection cooperation' (n 145).

[148] European Parliament, 'Report on the proposal for a directive of the European Parliament and of the Council amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on consumer protection cooperation ' (2008) A6-0318/2008, 163-164.

[149] ibid 88-89.

[150] Article 29 Data Protection Working Party, 'Opinion 1/2009 on the proposals amending Directive 2002/58/EC on privacy and electronic communications (e-Privacy Directive)' (2009) 00350/09/EN WP 159, 10.

[151] Commission, 'Amended proposal for a Directive of the European Parliament and of the Council Amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sectors and Regulation (EC) No 2006/2004 on consumer protection cooperation' (2008) COM(2008)723 final.

[152] Council, 'Common position adopted by the Council on 16 February 2009 with a view to the adoption of a Directive of the European Parliament and of the Council amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on consumer protection cooperation' (2009) 16497/1/08 REV 1.

settings.[153] Under this version, cookies would be allowed only with 'prior consent, which may be given by way of using the appropriate settings of a browser or another application'.[154] Again, certain compromises were made when the final version was adopted, with the Parliament's initial amendments somehow watered down. First, the explicit wording 'prior' that qualifies consent was removed. Second, a new recital was added to the reform package, stating:

> 'It is therefore of paramount importance that users be provided with clear and comprehensive *information* […]. The methods of providing *information* and offering the *right to refuse* should be as user-friendly as possible. Exceptions to the obligation to provide *information* and offer the *right to refuse* should be limited […].'[155]

The repetitive phraseology of 'information' and 'right to refuse', as will be shown below, is interpreted by some as an indication of retaining the same regulatory model. Third, while the reference to browser settings remains part of the reform, it was moved to Recital 66 instead of staying in the main-text provision. It reads: 'Where it is technically possible and effective, in accordance with the relevant provisions of [the DPD], the user's consent to processing may be expressed by using the appropriate settings of a browser or

---

[153] European Parliament, 'Recommendation for Second Reading on the Council common position for adopting a directive of the European Parliament and of the Council amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities for the enforcement of consumer protection laws' (2009) A6-0257/2009.

[154] ibid 56.

[155] Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws [2009] OJ L337/11, Recital 66 (emphasis added).

other application.' In practice, such changes have caused confusions and controversies over how the revised Article 5(3) should be implemented. Some hope of clarification has been pinned to the ongoing reform of the ePrivacy framework, with a Regulation proposed to replace the Directive.[156]

(b) The rise and fall of browser-based approaches

Since the amendment of the ePrivacy Directive in 2009, it has been widely discussed what would count as valid consent under the new Article 5(3). The particular strand of debates concerning browser settings became dominant among the stakeholders, not too surprisingly given the explicit mentioning of this approach in the 'cookie law'. However, there has been no consensus on how browser behaviour as well as how the action or inaction by a user should be interpreted. This concerns two technical solutions that rely on slightly different sets of actors in the industry.

Browser default settings

Since cookies are completely dependent on the functioning of a web browser, the browser itself may, technically speaking, have full control over which (categories of) cookies to allow and how such cookies are stored and accessed. Almost all modern web browsers allow users to enable or disable cookies, with a greater or lesser degree of customisability and varying configurations of default settings. The problem is, if a browser is set to have cookies enabled, does that amount to consent for the use of cookies? What if cookies are enabled by default? Does it make a difference if they are enabled manually? While trade groups representing the OBA industry believe that

[156] Commission, 'Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications)' (2017) COM(2017) 10 final.

Recital 66 means (default) browser settings can be seen as sufficient consent,[157] this is clearly not the position taken by regulators and researchers.[158] For example, the Working Party has consistently rejected the idea that default settings may serve as consent for cookies.[159] On national level, the UK Government as well as the ICO are both of the view that default browser settings would not fulfil the requirement of consent under the amended ePrivacy Directive.[160]

Since browser settings play a key role in allowing or blocking the use of cookies and default settings have a significant 'nudge' effect on user decisions, it is suggested that the responsibility to allow users to signify meaningful consent for cookies should fall upon browser manufacturers.[161] In fact, the UK

---

[157] Internet Advertising Bureau UK, *BIS Consultation on Implementing the Revised EU Electronic Communications Framework: IAB UK Response* (2010) <https://www.iabuk.net/sites/default/files/IABUKresponsetoBISconsultationonimplementingtherevisedEUElectronicCommunicationsFramework_7427_0.pdf> accessed 18 January 2018; Out-Law.com, 'Advertisers Say That New Cookie Law Is Met by Browser Settings' *Out-Lawcom* (24 November 2009) <https://www.out-law.com/page-10550> accessed 18 January 2018; Interactive Advertising Bureau, 'Telecom Package: Publishers and Online Marketers Welcome New Provisions on Cookies' (2009) <http://www.iab.it/iab-news/telecom-package-publishers-and-online-marketers-welcome-new-provisions-on-cookies/> accessed 18 January 2018.

[158] See McStay, 'I Consent: An Analysis of the Cookie Directive and Its Implications for UK Behavioral Advertising' (n 88); Kosta, 'Peeking into the Cookie Jar' (n 140).

[159] Article 29 Data Protection Working Party, 'Opinion 2/2010 on online behavioural advertising' (n 47) 14; Article 29 Data Protection Working Party, 'Opinion 15/2011 on the definition of consent' (n 86) 35.

[160] Media and Sport Department for Culture, *Open Letter on the UK Implementation of Article 5(3) of the e-Privacy Directive on Cookies* (2011) 5 <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/77638/cookies_open_letter.pdf> accessed 18 January 2018.

[161] See Article 29 Data Protection Working Party, 'Recommendation 1/99 on Invisible and Automatic Processing of Personal Data on the Internet Performed by Software and Hardware' (1999) 5093/98/EN/final WP 17, 7; Article 29 Data Protection Working Party, 'Opinion 2/2010 on online behavioural advertising' (n 47) 13-15; European Data Protection Supervisor, *Opinion of the European Data Protection Supervisor on Promoting Trust in the Information Society by Fostering Data Protection and Privacy* (2010) 18 <https://edps.europa.eu/sites/edp/files/publication/10-03-19_trust_information_society_en.pdf> accessed 18 January 2018; Ignacio N. Cofone, 'The Way the Cookie Crumbles: Online Tracking Meets

Government has pledged 'to continue to work with browser manufacturers to see if browsers can be enhanced to meet the requirements of the revised Directive.'[162] However, such efforts have failed to come to any substantial fruition, whether on national or EU level. This approach seems doomed to failure for several reasons. Politically, the EU or its Member States do not appear to have much leverage over mainstream browser manufacturers who are all based in the US[163] and whose business models rely heavily on online tracking.[164] Technically, even if web browsers have been improved with sufficient sophistication, it would still be impossible to make sure all users have updated their browsers to the latest version, if they use those major browsers at all.[165] Legally, it is questionable whether browser-level consent — even assuming this is done by changing the default settings on a voluntary, informed basis — can be considered 'specific' or 'ambiguous',[166] as such settings can essentially be considered blanket permissions to future use of cookies, although it would of course depend on the specific design.

Do Not Track ('DNT')

Even if such shortcomings could be overcome, endeavours on cookie settings would still suffer another fatal defect: Such solutions are technology-specific, meaning that they can only control cookies, but not other forms of

---

Behavioural Economics' [2017] International Journal of Law and Information Technology, 55-57.

[162] Department for Culture (n 160) 2.

[163] Cofone (n 161) 55.

[164] Brian Palmer, 'Why Is Microsoft Fighting So Hard Over Internet Explorer?' *Slate* (17 December 2009) <http://www.slate.com/articles/news_and_politics/explainer/2009/12/why_is_microsoft_fighting_so_hard_over_internet_explorer.html> accessed 18 January 2018; Zuiderveen Borgesius, *Improving Privacy Protection in the Area of Behavioural Targeting* (n 77) 249.

[165] Information Commissioner's Office, 'Guidance on the rules of cookies and similar technologies' (n 133) 15.

[166] Article 29 Data Protection Working Party, 'Opinion 2/2010 on online behavioural advertising' (n 47) 14; N. van Eijk and others, 'Online Tracking: Questioning the Power of Informed Consent' (2012) 14(5) Info 57, 60.

tracking device, such as Flash Cookies or device fingerprinting.[167] In this regard, an alternative browser-based approach was proposed to cover all sorts of online tracking. In a 2010 report, the US FTC suggested that '[t]he most practical method of providing uniform choice for online behavioral advertising would likely involve placing a setting similar to a persistent cookie on a consumer's browser and conveying that setting to sites that the browser visits, to signal whether or not the consumer wants to be tracked or receive targeted advertisements.'[168] This has developed into what is now known as Do Not Track ('DNT'). Unlike browser settings, switching on DNT itself does not block access to cookies, but simply sends out a request of non-tracking to the websites. It would technically be up to the websites how this request should be interpreted and complied with. In fact, when Microsoft decided to have DNT enabled by default on Internet Explorer 10 in 2012, it was immediately met with massive criticisms from the advertising industry, who threatened to ignore such DNT flags.[169] As such, for DNT to constitute a meaningful choice that is effectively honoured, it must be enforceable either by legislation or by industrial self-regulation.

However, little consensus has been achieved on how DNT should be translated into practical measures. On the legislative front, while a number of bills have been filed in the US Congress, none has been enacted.[170] In the EU,

---

[167] See Article 29 Data Protection Working Party, 'Opinion 9/2014 on the application of Directive 2002/58/EC to device fingerprinting' (n 70).

[168] Federal Trade Commission, 'Protecting Consumer Privacy in an Era of Rapid Change: A Proposed Framework for Businesses and Policymakers (Preliminary FTC Staff Report)' (2010), 66.

[169] Joshua A.T. Fairfield, 'Do-Not-Track as Default' (2013) 11(7) Northwestern Journal of Technology and Intellectual Property 575, 578-579.

[170] Library of Congress, 'Congress.gov' <https://www.congress.gov/advanced-search/legislation?enterTerms=%22Do+Not+Track%22&search=search> accessed 18 January 2018. See also Dawn Chmielewski, 'How 'Do Not Track' Ended Up Going Nowhere' *Recode* (4 January 2016) <https://www.recode.net/2016/1/4/11588418/how-do-not-track-ended-up-going-nowhere> accessed 18 January 2018. For state-level initiatives in the US, see Irene

Article 5(3) and Recital 66 of the revised ePrivacy Directive have remained the only relevant provisions. On both sides of the Atlantic, much hope has been pinned on industry-wide standardisation of the DNT mechanism. Since 2011, the World Wide Web Consortium (W3C) — an international organisation that develops standards for the Web — has been working with representatives of various stakeholders on 'defining mechanisms for expressing user preferences around Web tracking'.[171] The standardisation work has been divided into two building blocks, Tracking Preference Expression (DNT) and Tracking Compliance and Scope (TCS), with the former dealing with the technical specifications of signal transmission[172] and the latter with interpretation and compliance of such signals.[173] At the time of writing, both standards are in the status of 'Candidate Recommendation' but not yet adopted.[174]

A significant divergence exists on how the signals of 'DNT:0' (preference to allow tracking) and 'DNT:1' (preference of non-tracking) should be honoured in practice. For instance, while the Working Party has constantly insisted that 'DNT:1' means 'Do Not Collect', 'Do Not Use' and 'Do Not Share' without exceptions,[175] this is clearly not the approach taken by the W3C in their

Kamara and Eleni Kosta, 'Do Not Track Initiatives: Regaining the Lost User Control' (2016) 6(4) International Data Privacy Law 276, 283-284.

[171] W3C, 'Tracking Protection Working Group' <https://www.w3.org/2011/tracking-protection/> accessed 18 January 2018.

[172] W3C, 'Tracking Preference Expression (DNT): W3C Candidate Recommendation 19 October 2017' (2017) <https://www.w3.org/TR/tracking-dnt/> accessed 18 January 2018.

[173] W3C, 'Tracking Compliance and Scope: W3C Candidate Recommendation 26 April 2016' (2016) <https://www.w3.org/TR/tracking-compliance/> accessed 18 January 2018.

[174] W3C, 'Tracking Protection Working Group' (n 171).

[175] Article 29 Data Protection Working Party, 'Opinion 04/2012 on Cookie Consent Exemption' (2012) 00879/12/EN WP 194, 10; Article 29 Data Protection Working Party, *Article 29 Data Protection Working Party comments in response to W3C's public consultation on the W3C Last Call Working Draft, 14 July 2015, Tracking Compliance and Scope* (2015) 2 <http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2015/20151001_letter_of_the_art_29_wp_w3c_compliance.pdf> accessed 18 January 2018.

recommended draft.[176] Also, in the case of 'DNT:0', the W3C suggests this should amount to 'no restrictions on collection or use of data' unless overridden by the website's own statement or consent obtained otherwise.[177] Yet, the Working Party has repeatedly emphasised that '[a] DNT:0 signal must not be interpreted by a data controller as consent for anything other than clearly defined tracking activities' and that in the absence of a clear DNT value, it should be assumed that 'a user is not aware of tracking'.[178] In short, the Working Party's position on DNT can be summarised as 'no means no, silence means no, and yes does not necessarily mean yes'. Therefore, a meaningful consensus on what DNT signals represent — one that would be both considered compliant by European regulators and accepted by various stakeholders — does not seem to have a bright future.

(c) 'By continuing to use this site' … can one dis*agree?*

Given the technical deficiency as well as the legal uncertainty of these solely-browser-settings-based solutions, it would be reasonable to expect website operators to continue to ask users for consent in traditional manners: banners, dialogues and walls. In any event, as mentioned above, browser settings and signals can be overruled by consent that is obtained in more specific ways.[179] In a Working Document, the Working Party provides certain guidelines on

---

[176] W3C, 'Tracking Compliance and Scope: W3C Candidate Recommendation 26 April 2016' (n 173) s 3.2.
[177] ibid s 3.1.
[178] Article 29 Data Protection Working Party, *Article 29 Data Protection Working Party comments in response to W3C's public consultation on the W3C Last Call Working Draft, 24 April 2014, Tracking Preference Expression (DNT)* (2014) 3 <http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2014/20140606_wp29_ts_standardisation_letter_to_w3c.pdf> accessed 18 January 2018; Article 29 Data Protection Working Party, *Article 29 Data Protection Working Party comments in response to W3C's public consultation on the W3C Last Call Working Draft, 14 July 2015, Tracking Compliance and Scope* (n 175) 3.
[179] Zuiderveen Borgesius, *Improving Privacy Protection in the Area of Behavioural Targeting* (n 77) 254.

obtaining consent for cookies.[180] Of course, the 'best practice' that is most likely compliant with the ePrivacy Directive would be 'a clear, comprehensive and visible notice' supported with sufficient information, [181] with no cookies set 'before that user has signalled their wishes'[182] 'by clicking on a button or link or by ticking a box',[183] choosing 'between the option to accept some or all cookies or to decline all or some cookies', with 'the possibility to continue browsing the website without receiving cookies'.[184] However, it is obvious that not all websites will be willing to adhere to such a highest standard, and a common approach is to display a banner with a link to more information on cookies, and a statement that 'by continuing to use this site, you agree to our use of cookies.' Such a practice raises two concerns about compatibility with the ePrivacy Directive: whether continued use of a service can be considered a valid form of consent, and whether the provision of service can be made conditional on such consent.

The controversy of 'implied consent'

The disputes over how Article 5(3) should be enforced in the context of online tracking result largely from the diverse understanding of what counts as consent under the ePrivacy Directive. In the UK, for example, this has caused public debates on whether the idea of 'implied consent' may be acceptable. In May 2011, the ICO published a notice on the implications of the revised Article 5(3), which stated that, to obtain meaningful consent, the data controller would 'need to provide information about cookies and *obtain consent before a cookie is set* for the first time', which requires 'a positive indication […]

[180] Article 29 Data Protection Working Party, 'Working Document 02/2013 providing guidance on obtaining consent for cookies' (2013) 1676/13/EN WP 208.
[181] ibid 3.
[182] ibid 4.
[183] ibid.
[184] ibid 5.

most commonly obtained by asking the user to tick a box' and that '[a]ny attempt to gain consent that relies on users' ignorance about what they are agreeing to is unlikely to be compliant.'[185] However, this interpretation was not welcome by the UK Government, who issued an open letter later in the same month, explaining its position that the definition of consent 'is not time bound — i.e. there is no constraint on when consent may be given', and that 'Article 5 of the revised e-Privacy Directive does not specify that the consent must be "prior consent".'[186] This led to a softened tone in the ICO's updated guidance in December 2011.[187] In that version, while the ICO acknowledged that it was hard to see how consent could be obtained after the cookies were set, it also recognised the difficulties for some websites to obtain prior consent.[188] Moreover, the concept of 'implied consent' was discussed in that guidance for the first time. It was suggested that '[a]t present evidence demonstrates that general awareness of the functions and uses of cookies is simply not high enough for websites to look to rely entirely in the first instance on implied consent.'[189] However, '[a]s consumer awareness increases over the next few years it may well be easier for organisations to rely on that shared understanding to a greater degree.'[190] In less than a year, the ICO seemed to have entirely changed to embrace the idea of 'implied consent' in a more

---

[185] Information Commissioner's Office, *Changes to the Rules on Using Cookies and Similar Technologies for Storing Information* (2011) 6-7 <https://web.archive.org/web/20110511051726/http:/www.ico.gov.uk/~/media/documents/library/Privacy_and_electronic/Practical_application/advice_on_the_new_cookies_regulations.pdf> accessed 18 January 2018 (emphasis added).

[186] Department for Culture (n 160) 2-3.

[187] Information Commissioner's Office, *Guidance on the Rules on Use of Cookies and Similar Technologies* (2011) <https://web.archive.org/web/20111213204833/http://www.ico.gov.uk/news/latest_news/2011/~/media/documents/library/Privacy_and_electronic/Practical_application/guidance_on_the_new_cookies_regulations.ashx> accessed 18 January 2018.

[188] ibid 6.

[189] ibid.

[190] ibid.

straightforward manner. Its renewed guidance in May 2012 considered implied consent 'always […] a reasonable proposition' and 'certainly a valid form of consent' that serves as 'an option that was perhaps more practical than the explicit opt-in model.'[191] To explain what would suffice as implied consent, the guidance goes on: 'This might for example be visiting a website, moving from one page to another or clicking on a particular button.'[192] This remains the ICO's position to date, reiterated in another guidance issued in 2016.[193]

The UK's position on consent drew support from 13 Member States, who argued that the change of wording in Article 5(3) was not intended to alter the existing requirement.[194] Instead, they believed that, based on Recital 66 — 'providing information and offering the right to refuse'[195] — the regime should remain an 'opt-out' one. However, this approach has received constant criticisms. As early as in 2011, the Working Party had stated that implied consent was incompatible with the definition of consent under the DPD.[196] The change of language in Article 5(3) clarifies that the so-called 'opt-out' consent is no longer a legitimate basis for the use of cookies.[197] This is also the interpretation endorsed by the Working Party in a 2013 Working Document,

---

[191] Information Commissioner's Office, 'Guidance on the rules of cookies and similar technologies' (n 133) 6-7.
[192] ibid 7.
[193] Information Commissioner's Office, 'Guide to the Privacy and Electronic Communications Regulations' (2016), 27.
[194] Kosta, 'Peeking into the Cookie Jar' (n 140) 391.
[195] Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws, Recital 66.
[196] Article 29 Data Protection Working Party, 'Opinion 15/2011 on the definition of consent' (n 86) 24.
[197] McStay, 'I Consent: An Analysis of the Cookie Directive and Its Implications for UK Behavioral Advertising' (n 88). See also Kosta, *Consent in European Data Protection Law* (n 7) 188-202; Cofone (n 161) 40-42.

where they maintain that 'consent should be sought before cookies are set or read' and 'through a positive action or other active behaviour'.[198] While the Working Party has clarified its position against the practice of implied consent in the form of 'by continuing to use this site'-style notice in the latest guidelines regarding the GDPR,[199] the lack of enforcement or legal actions on this matter[200] means that the effectiveness of this interpretation would depend on the final adopted text of the ePrivacy Regulation, as well as the enforcement thereof.

The tracking walls

Another issue concerning the way consent is obtained for cookies has to do with the 'freely given' element as discussed above. It is subject to constant debate whether online service providers can or should be allowed to ask users for consent for tracking as a condition for access to their services. Practices of denying service to those who refuse to consent are known as 'tracking walls' or 'cookie walls'.[201] Recital 25 of the ePrivacy Directive provides that '[a]ccess to specific website content may still be made conditional on the well-informed acceptance of a cookie or similar device, if it is used for a legitimate purpose.'[202] This, at first glance, seems to suggest that tracking walls are not prohibited. However, the Working Party has taken a more nuanced reading of this recital, explaining that '[t]he emphasis on "specific website content" clarifies that websites *should not* make conditional "general access" to the site on acceptance

[198] Article 29 Data Protection Working Party, 'Working Document 02/2013 providing guidance on obtaining consent for cookies' (n 180) 4.
[199] Article 29 Data Protection Working Party, 'Guidelines on consent under Regulation 2016/679' (n 76) 17.
[200] J. A. Luzak, 'Privacy Notice for Dummies? Towards European Guidelines on How to Give 'Clear and Comprehensive Information' on the Cookies' Use in Order to Protect the Internet Users' Right to Online Privacy' (2014) 37 Journal of Consumer Policy 547, 556.
[201] See Zuiderveen Borgesius and others (n 101).
[202] ePrivacy Directive, Recital 25.

of all cookies but can only limit certain content if the user does not consent to cookies'.[203] In other words, a user's refusal to consent on tracking cookies for, say, advertising purpose may only prevent the user from using advertising-related functions of the website, not all of it. Such an interpretation would sound more reasonable with in mind Article 7(4) of the GDPR, under which consent for data processing not necessary for the service provision would be presumed not 'freely given' if the service is made conditional on such consent. To minimise legal uncertainty and fragmented interpretation, however, further clarification on this point would still be needed.

(d) The proposed ePrivacy Regulation

As soon as the GDPR was adopted in 2016, discussions on updating the ePrivacy Directive so as to put it in line with the new standards set out by the GDPR had begun. In early 2017, the Commission proposed a draft ePrivacy Regulation, intended to replace the ePrivacy Directive.[204] At the time of writing, an amended proposal has been drafted by a Committee of the Parliament and is pending for approval in first reading. The Council is also working in parallel on its own amended version. Article 8(1) of the Commission's proposal can be seen as an updated continuation of Article 5(3) of the ePrivacy Directive. While the regulatory regime under the proposed provision remains by and large similar to the old one, two significant changes should be noted. First, the definition of consent is to be completely aligned with the GDPR; second, as suggested by the Working Party,[205] first-party web

[203] Article 29 Data Protection Working Party, 'Working Document 02/2013 providing guidance on obtaining consent for cookies' (n 180) 5 (emphasis original).
[204] Commission, 'Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications)' (n 156).
[205] Article 29 Data Protection Working Party, 'Opinion 04/2012 on Cookie Consent Exemption' (n 175) 11.

audience measuring is to be introduced as a new ground for the use of cookies.[206] However, the practice of using 'by continuing to use this site' banners remains unaddressed. The Parliament version, along with certain minor amendments to the proposal, attempts to introduce two more exemptions — necessity to ensure security of the device, and necessity for executing an employee's tasks — which have little to do with tracking for OBA purposes. Much more relevant is a newly proposed article, which reads: 'No user shall be denied access to any information society service or functionality […] on grounds that he or she has not given his or her consent under Article 8(1)(b) to the processing […] that is not necessary for the provision of that service or functionality.'[207] This provision, if eventually adopted by the legislators as it is, would be a final blow to the current practices of tracking walls. Having said that, as clearly demonstrated in the course of legislating the GDPR, it is not uncommon for the Parliament to make certain last-minute compromises and accept some weakened measures. However, given the failure of the browser-based approaches to consent, as well as the new risks of the ubiquitous tracking practices flagged up throughout this study, the stakes involved in such compromises this time would be greater than ever before.

4.5 Data protection principles: An abstract 'safety net'

Identifying the appropriate legal basis for processing marks only the first hurdle in establishing the lawfulness of data processing. Even more complicated is to prove that the processing in line with a set of data protection

[206] Commission, 'Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications)' (n 156) arts 8(1)(b) & (d).

[207] European Parliament, 'Report on the proposal for a regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications)' (2017) A8-0324/2017, 58.

principles, which are abstract, sophisticated, inter-related and organic — and they are all of paramount importance in a data protection regime. Among the seven principles expressly enumerated in the GDPR,[208] each and every one is relevant to our study on OBA, or big data in more general terms. However, these principles can each be a topic for an independent study, and it is not the task to conduct a comprehensive analysis here. Instead, it would make more sense to focus on the three particularly relevant ones: lawfulness, fairness and transparency; data minimisation; and purpose limitation. Most of these principles are more or less already in place in the DPD, but certain subtle changes may make a significant difference.

4.5.1 Lawfulness, fairness and transparency

It has been long established as a principle in the DPD that personal data must be processed fairly and lawfully. Now a third element of transparency is added to this principle. There is little doubt that lawfulness, fairness and transparency are all the most fundamental principles in data protection law. However, as with many of the legal principles, these three terms are arguably of more declarative than practical importance. Their overarching status gives them an almost all-encompassing, yet highly ambiguous scope. Having said that, from a practical point of view, in particular with the case study of OBA, certain specific implications of these principles can be seen more clearly, particularly based on the discussion above on the legal bases provided by Article 6.

To begin with, the lawfulness requirement can be largely met by ensuring an appropriate legal ground for the processing. Recital 40 of the GDPR makes it quite clear that '[i]n order for processing to be lawful, personal data should be processed on the basis of the consent of the data subject concerned or some

[208] GDPR, art 5.

other legitimate basis'.[209] As concluded in the previous section, the only appropriate (and practical) legal ground for data processing in an OBA context would be consent. When it comes to sensitive data, this needs to be *explicit* consent. It is also arguable that having a lawful basis for the processing in question is only a necessary but not sufficient condition for the lawfulness principle. For example, this principle may also suggest that data processing must observe the rules of other areas of law, such as contract law or consumer protection law.

In this regard, the principle of fairness is perhaps a better place to discuss the influence from other branches of EU law on data protection, as the concept of '(un)fairness' has a longer history in, *inter alia*, contract law or consumer protection law. The Unfair Contract Terms Directive defines an unfair contract term as one that 'has not been individually negotiated [… and …] causes a significant imbalance in the parties' rights and obligations arising under the contract, to the detriment of the consumer.'[210] Translating this idea to data protection law would mean that data controllers must not exploit their advantageous position and in effect impose a requirement on the data subject. This would in theory involve an assessment of whether there exists an imbalance of power between the data controller and the data subject.[211] This concerns the discussion about the 'freely given' element of consent, and in particular in the context of OBA, the conditionality test set out in Article 7(4). In other words, the practice of bundling up consent for unnecessary data processing with access to service is assumed exploitative by the GDPR, and

[209] ibid Recital 40.
[210] Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts [1993] OJ L95/29 ('UCTD'), art 3(1).
[211] For a similar theory that examines the fairness principle with a balancing approach in the case of power asymmetries, see Damian Clifford and Jef Ausloos, 'Data Protection and the Role of Fairness' (2017) 29/2017 CiTiP Working Paper Series <https://ssrn.com/abstract=3013139> accessed 14 December 2017.

therefore, presumably, cannot be compatible with the fairness principle. Another interpretation of fairness may have more to do with fair treatment to individuals. The Working Party is of the view that use of data 'may be unfair and create discrimination, for example by denying people access to employment opportunities, credit or insurance, or targeting them with excessively risky or costly financial products.' [212] According to this understanding, data use for OBA purposes should not create any discriminatory effects.

The third principle, transparency, is in close connection with the 'specific' and 'informed' requirements of consent when it comes to OBA, since, again, consent serves as the only practical lawful basis. Of course, as a general principle, transparent information is required regardless of the legal ground chosen and throughout the entire lifecycle of data processing. That means, OBA operators are obliged to provide information not merely at the point of obtaining consent, but throughout the entire lifecycle of data processing. Circumstances under which additional information may be required to fulfil the transparency requirement include, *inter alia*, the receipt of and decision on data subject requests,[213] further processing of personal data,[214] rectification or erasure of personal data,[215] and personal data breaches.[216]

4.5.2 Purpose limitation: Primary and secondary uses

The principle of purpose limitation in the GDPR replicates nearly word-by-word its equivalent in the DPD. This principle requires personal data to be

[212] 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 251rev.01, 10.
[213] GDPR, arts 12(3) & (4).
[214] ibid art 13(3).
[215] ibid art 19.
[216] ibid art 34.

'collected for specified, explicit and legitimate purpose and not further processed in a manner that is incompatible with those purposes'.[217] This principle holds particular significance to the case of OBA, as it is constantly questioned what amounts to compatible secondary use when personal data is used for marketing purposes. Here, two separate issues could be concerned: a) Can OBA constitute a compatible secondary use if the personal data in question was originally collected for a different purpose of, say, provision of online services? b) If personal data is processed for the (parallel) primary purpose of OBA, what types of further use will be allowed as compatible further processing? Before these two questions can be answered, it is necessary to have an overview of the key elements of this principle.

In examining the essence of the principle, the Working Party breaks purpose limitation into two components: purpose specification and compatible use.[218] The first building block, purpose specification, stipulates that personal data must be 'collected for specified, explicit and legitimate purposes'.[219] These three elements have been more or less translated into the legal bases for lawful processing. For consent to be valid, it must be *specific* and *explicit* in terms of the content of the consent, including the purposes for which the processing is consented. The *legitimacy* condition, albeit not expressly included as part of the definition or perquisites for valid consent, can be inferred from the nature of consent. The Working Party takes the view that the legitimacy element in the purpose specification principle 'also requires that the purposes must be in accordance with all provisions of applicable data protection law, as well as other applicable laws such as employment law,

[217] ibid art 5(1)(b).
[218] Article 29 Data Protection Working Party, 'Opinion 03/2013 on purpose limitation' (n 84).
[219] GDPR, art 5(1)(b).

contract law, consumer protection law, and so on.'[220] As we have analysed previously, consent is by definition a mutual agreement, and therefore its validity is also subject to other sectors of law that might lay down further restrictions on contractual arrangements. This way, the examination of legitimacy as required by purpose specification forms an intrinsic part of identifying valid consent. Considering that consent is very likely the only legal basis for the OBA industry to rely on, the fulfilment of the conditions set down by provisions regarding consent may, generally speaking, suffice to satisfy the requirement of purpose specification.

The more problematic part of purpose limitation lies in its second building block — compatible use. This component requires that personal data not be further processed in a manner that is incompatible with those purposes as specified at the time of primary use. The implications of this are two-fold. On the one hand, further use of personal data for incompatible purposes is prohibited; on the other, compatible use of such data should be allowed *with no need for any legitimising basis*. At the time of the DPD, the Working Party is of the view that, since 'compatibility' and 'lawful ground' are two independent, accumulative requirements, further processing even for compatible purposes must also be based on one of the six lawful grounds.[221] While the Working Party seems to retain the same interpretation in its latest Guidelines,[222] this is clearly not the interpretation taken by the GDPR. Recital 50 clarifies that '[i]n such a case [where compatible use is established], no legal basis separate from that which allowed the collection of the personal data is

[220] Article 29 Data Protection Working Party, 'Opinion 03/2013 on purpose limitation' (n 84) 20.
[221] ibid 11.
[222] Article 29 Data Protection Working Party, 'Guidelines on consent under Regulation 2016/679' (n 76) 12.

required.'[223] It would make more sense to follow the GDPR's own approach, because were a further legal basis required for further processing, it would be pointless to distinguish 'initial processing' and 'further processing' — all 'further processing' justified by its own lawful basis would be *de facto* turned into 'initial processing'. This change — or rather, clarification — will have profound implications, at least in theory, in that it has essentially transformed the nature of purpose limitation from a purely prohibitive principle (incompatible use is prohibited) to a partly permissive one (compatible use is permitted with no need for an additional legal basis). That said, the impact in practice might be neutralised by the fact that the compatibility test is quite high a threshold itself.

When assessing whether the intended further processing is compatible with the original purpose, Article 6(4) provides a list of factors that should be taken into account: the link between the original and new purposes; the context of data collection, in particular the relationship between data subjects and the data controller; the nature of personal data; the possible consequences for data subjects; and appropriate safeguards.[224] These criteria are helpful for the two questions regarding OBA raised in the beginning of this section: (a) Is OBA a compatible secondary purpose? (b) If OBA is the primary purpose, what are the secondary purposes compatible with it?

For the first question, the Working Party has constantly reiterated that online marketing cannot stand as a compatible purpose if the initial purpose is, for instance, the provision of online services. By way of example, the Working Party illustrates in a 2013 Opinion that for a photo-sharing social media service, advertising is not a compatible purpose with the original one

[223] GDPR, Recital 50.
[224] ibid art 6(4).

(allowing users to share photos) for a number of reasons: The two purposes are 'clearly unrelated'; the nature of the data can be sensitive; it may cause self-censorship; and there is an imbalance of power between the website and its users.[225] This is further confirmed in the latest Guidelines with regard to the GDPR, with another example of incompatibility between the purpose of suggesting new movies by a cable TV network and that of showing targeted advertising.[226] It is therefore the view of the Working Party that OBA would not count as a compatible purpose. Accordingly, repurposing existing data sets for OBA would need to be justified by one of the lawful bases, presumably consent.

Much more complicated is the second issue, namely the scope of compatible uses when personal data is collected for marketing purposes. Indeed, most online services today have explicitly listed personalised content, including advertisements, as one of their purposes of data collection. Google, for example, states that user data is used to 'offer you tailored content — like giving you more relevant search results and ads'.[227] Of course, as explained above, the mere reference to 'relevant ads' alone would not suffice to justify all three complex stages of OBA: tracking, profiling and targeting. More details are required about how data is actually handled during each of these phases.

Assuming all information regarding the originally envisaged purposes and processing operations has been sufficiently covered, the next question would be how newly developed OBA techniques may rely on the original legal basis. In the tracking stage, for instance, the ad network provider might decide to

---

[225] Article 29 Data Protection Working Party, 'Opinion 03/2013 on purpose limitation' (n 84) 60-61.

[226] Article 29 Data Protection Working Party, 'Guidelines on consent under Regulation 2016/679' (n 76) 12.

[227] Google, 'Privacy Policy' (2017) <https://static.googleusercontent.com/media/www.google.co.uk/en/uk/intl/en_uk/policies/privacy/google_privacy_policy_en_uk.pdf> accessed 26 September 2017.

expand the tracking device from first-party to also third-party cookies, or from cookie-based techniques to non-cookie-based device fingerprinting, or from certain categories of data to more categories. These practices would lead to a higher level of accuracy for the tracking device, which on the one hand can be regarded improvements that would better serve the OBA purpose, but on the other also represents an escalating level of intrusiveness. Depending on the nature of data and the potential impact on the data subject, such enhanced techniques might well be found incompatible with the initial ones.

A similar issue arises from certain practices in the profiling stage as well. For instance, an OBA operator may decide to merge user profiles across services to build up a fuller profile, which was, as mentioned in Chapter 1, precisely what Google did in 2012 with its updated privacy policy and what Facebook did in 2016 with its plan to aggregate data from WhatsApp. The Working Party already made it clear in 2000 that '[i]f ad network providers want to use information gathered for behavioural advertisement for secondary, *incompatible purposes, for example across services*, they need additional legal grounds to do so'.[228] Yet, in subsequent enforcement actions against Google and Facebook, neither the Working Party nor national authorities have explicitly pointed out that such aggregation of data across services violates the principle of purpose limitation.[229] Nevertheless, it should remain clear that for

---

[228] Article 29 Data Protection Working Party, 'Opinion 2/2010 on online behavioural advertising' (n 47) 20 (emphasis added).

[229] For a full coverage of the actions taken against Google, see Rauhofer, 'Of Men and Mice' (n 87). For actions against Facebook, see Article 29 Data Protection Working Party, 'Letter of 16 December 2016' (2016) <https://ec.europa.eu/newsroom/document.cfm?doc_id=40927> accessed 21 February 2018; Article 29 Data Protection Working Party, 'Letter of 24 October 2017' (2017) <https://ec.europa.eu/newsroom/just/document.cfm?doc_id=47964> accessed 21 February 2018; The Hamburg Commissioner for Data Protection and Freedom of Information, 'Administrative Order Against the Mass Synchronisation of Data Between Facebook and WhatsApp' (2016) <https://www.datenschutz-hamburg.de/fileadmin/user_upload/documents/Press_Release_2016-09-27_Adminstrative_Order_Facebook_WhatsApp.pdf> accessed 21 February 2018; CNIL,

the purpose of OBA, combining user profiles from multiple sources can hardly be held compatible with such original purposes.

In the stage of targeting, the issue of compatibility is particularly complicated in view of real-time bidding. It remains highly uncertain whether it would be allowed as a compatible further use of data if an OBA operator switches from a non-RTB-based solution to an RTB-based one. It is however argued here that adopting an RTB-based solution cannot be considered a compatible use if this is not initially envisaged, for the following reasons: a) real-time bidding in itself is designed to maximise the profit from a successful match, although it does improves the relevance of the ads as well; b) the lack of direct connections between the user and potential advertisers or bidding agents would not give the user any reasonable expectation that their personal data is shared with these third parties; and c) the benefits of adopting certain safeguards are largely reversed by such new techniques. For example, the existence of cookie matching largely cancels out the positive effects of pseudonymisation. While it is true that the user IDs communicated between the ad network provider and the bidders are always encrypted pseudonyms, the practice of cookie matching may potentially enable the bidder to identify the user in their own databases and trace the user's online footprints across the Internet. As such, to ensure the targeting phase does not contravene the purpose limitation principle, OBA operators would need to either obtain valid consent for activities related to real-time bidding (so as to turn it into a primary use), or resort to an alternative technical model that shows a higher degree of compatibility.

'Data Transfer from WHATSAPP to FACEBOOK: CNIL Publicly Serves Formal Notice for Lack of Legal Basis' (2017) <https://www.cnil.fr/en/data-transfer-whatsapp-facebook-cnil-publicly-serves-formal-notice-lack-legal-basis> accessed 21 February 2018.

Apart from upgrading the tracking, profiling and targeting techniques, further processing can also take the form of analysing OBA data for new purposes. For example, OBA data collected initially for commercial targeting can be easily put into use for political targeting. Also, such data can be used to personalise or prioritise news feed or social media content. The risks of such instances of further use of OBA data have been underlined in Chapter 3 and should not be underestimated. It would be hard for the data controller to argue that such secondary purposes are compatible with that of OBA. When it comes to marketing, the Working Party takes a particularly strict test of compatibility. In an example concerning mobile apps designed to help users find nearby restaurants offering discounts (primary purpose), the use of location data for advertising pizza delivery (secondary purpose) is considered incompatible.[230] It is therefore hard to see how this type of repurposing of OBA data could be considered compatible with the purpose limitation principle.

4.5.3 Data minimisation: What is 'necessary'?

Another principle that deserves a closer examination is data minimisation. In fact, this principle has a scope broader than its name implies. Three particular criteria are set forth in Article 5(c): 'adequate, relevant and limited to what is necessary'. The spirit of this principle shares a lot in common with that of the purpose limitation principle in that all the three tests mentioned above should be applied with particular reference to the purposes of the processing. Neither the DPD nor the GDPR further defines what 'adequate' or 'relevant' means, but it is conceivable that in the cases where the personal data cannot serve the purposes for which the processing is intended, such criteria are not fulfilled. The third standard, namely 'limited to what is necessary', seems to fit the name

[230] Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (n 212) 11.

of the principle 'data minimisation' most and attracts most critical attention. In practice, OBA operators need to answer the question as to whether the data processing they carry out is indeed necessary for the purposes they claim.

The exact wording in the GDPR ('limited to what is necessary') is obviously stronger than that in the DPD ('not excessive'). This seems to be a compromise between an even stricter version in the Commission's proposal ('limited to the minimum necessary') and the Council's preliminary amendment (changing back to 'not excessive'). As explained in the statement of the Council's reasons for the final version, 'the principle of "data minimisation" has been adjusted to take into account the digital reality and with a view to establishing a balance between protection of personal data, on the one hand, and possibilities for controllers to process data, on the other hand.'[231] None of these documents has further clarified the practical implications of such a change, but it is clear that the legislators wish to see a standard of data protection higher than what it is now yet not too high for reasonable uses of personal data.

Despite the lack of substantial specifics in the text of legislation, much can be drawn from the previous discussion of related principles and legal grounds. The strongest connection can be identified with purpose limitation, as the assessment of whether certain processing exceeds what is necessary depends greatly on what the purpose of processing is supposed to be. Without a well-defined scope of the purpose, there would be no benchmark to determine necessity or the lack of it. To such an extent, purpose limitation underpins data minimisation. If the primary purposes are sufficiently specified, explicit and legitimate, and the secondary purposes are compatible with the primary ones,

[231] Council, 'Position of the Council at first reading with a view to the adoption of a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL 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) - Statement of the Council's reasons' (2016) ST 5419 2016 REV 1 ADD 1, 8.

then the next step would be to examine if the intended processing is actually *necessary* for those specific purposes.

The necessity test, as the CJEU interprets in *Schecke* in a data protection context, requires the processing in question to be 'strictly necessary', ensuring the measures adopted are consistent with the objective of the processing while at the same time cause less interference to the data subject.[232] Now that direct marketing is confirmed as a legitimate interest,[233] are tracking, profiling and targeting for OBA necessary for direct marketing? If OBA is indeed a necessary manner, are the techniques employed and the data collected by an ad network provider as shown in Chapter 1 necessary for OBA? With a sceptical tone, it can even be argued that OBA itself is not necessary at all, because direct marketing can be done through less intrusive ways, such as contextual advertising, although that would cost the efficiency of targeting.[234] Neither is the 'financial necessity' a convincing argument here, as proved in Chapter 2.

The popular practices among the OBA industry make it even harder to pass the necessity test. The techniques employed in the targeting phase, for example, are not merely intended to predict what might interest a particular user, they also enable the ad network provider to identify the advertisers who could afford more on reaching that user. There is no denying that in the offline world, it is commonplace for a middleman to work out a potential match from the supply and demand sides, such as a letting agent. In the setting of OBA, however, the data subject has no interest in, or expectation of, being picked by the one who wants to talk the most. At the end of day, unlike a real-*life* bidding, the money coming from the real-*time* bidding does not go to the seller, i.e. the

---

[232] Joined Cases C-92/09 and C-93/09 *Volker und Markus Schecke and Eifert* [2010] OJ C 13/6 para 77.
[233] GDPR, Recital 47.
[234] Zuiderveen Borgesius, 'Personal Data Processing for Behavioural Targeting: Which Legal Basis?' (n 77) 168.

Internet user, at least not in monetary form, but wholly to the auctioneer. Again, while Internet users indeed benefit from such 'free services', it can be argued that the degree of benefits is severely disproportionate to the potential risks, as well as the interests gained by the industry. As such, a strong case can be made against the current practices of the industry in the light of the data minimisation principle under the GDPR, unless the purpose is clearly defined to reflect the nature of real-time bidding and measures to achieve such a purpose is made transparent to the users.

4.6 The (new?) role of 'profiling' in the GDPR

4.6.1 The concept of profiling

Besides the principles and legal bases that OBA operators need to comply with, there is also a provision in the GDPR that holds particular relevance. 'Profiling' as a legal concept has been introduced into the GDPR for the first time in the history of EU data protection law. Article 4(4) gives the definition of profiling as 'any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements'.[235] A few key points can be quickly identified from the text of this provision.[236] First, profiling has to be 'automated processing', which means, practices involving personal profiles established only manually would not be considered as profiling. Second, the purpose of such processing has to be 'to evaluate certain personal aspects'. It follows that building up records of merely factual personal data

[235] GDPR, art 4(4).

[236] See also Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (n 212) 6-7.

without the intention to evaluate the data subjects is not 'profiling' under such a definition. Third, and perhaps most importantly, profiling is carried out over *personal data*, and therefore, the use of fully anonymous data for the purpose of evaluating an entirely unidentifiable natural person is not profiling as such. This thus leads to a somewhat interesting conclusion that, the express inclusion of profiling as a legal concept in the GDPR has not expanded the scope of data protection law; it has simply made it clearer. This distinction has not just theoretic merits, but also matters in practice, in particular to OBA-related activities.

It is almost indisputable that the construction of user profiles by an ad network provider is conducted by automatic means and intended for evaluating the users' personal aspects, namely 'personal preferences' or 'interests'. When it comes to the third point — whether the data concerned is personal data — however, the provider might make an argument, not least in the cases where the user's profile is not associated with their registered account, that the profiling does not involve any 'processing of personal data'. Of course, according to the analysis in Section 4.3, such data should generally be considered personal data because it is linked to unique cookie IDs. Here, what needs to be emphasised is that the new concept of 'profiling' does not change the boundaries of 'processing' or 'personal data'. Instead, it identifies a new subset of processing, removing a degree of uncertainty from the notion. In other words, it becomes clearer under the GDPR that the practices of OBA profiling are regulated by data protection law, but that change has more to do with the better-defined scope of personal data than with the introduction of the 'profiling' notion itself. In fact, the GDPR does not create any obligations or rights that are solely connected to profiling practices. The term appears in the text mostly in the phrase 'including profiling' as a means of emphasis,

clarification or exemplification[237] — although profiling is nevertheless one form of processing of personal data, and is therefore regulated by data protection law.[238]

4.6.2 Profiling and automated decision-making

Maybe the greater significance of defining 'profiling' lies in its relationship with the old concept of 'automated decision-making' retained from the DPD. Under Article 22 ('Automated individual decision-making, including profiling'), data subjects have the 'right not to be subject to a decision based solely on automated processing'. The concepts of 'profiling' and 'automated individual decision-making' are overlapping but do not necessarily take place simultaneously.[239] When profiling is employed in the context of automated individual decision-making, Article 22 applies; otherwise, the general rules of the GDPR apply.[240] It should be noted that although Article 22 is phrased as 'the right not to be subject to a decision […]', the Working Party consider this provision as a general prohibition rather than a right to be invoked.[241] Recital 71 is cited to supported this interpretation, which reads 'decision-making based on such processing, including profiling, should be allowed where [one of the exceptions applies]'. Three such exceptions are provided by Article 22(2). The first exception is substantially identical to one of the general grounds for data processing, 'performance of contract'. The conclusion above thus holds true here: It is very unlikely for an OBA marketer to rely on this basis

[237] For example, see GDPR, arts 13(2)(f), 14(2)(g), 15(1)(h), 21, 22, 35(3)(a), 47(1)(e).
[238] Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (n 212) 8-9.
[239] ibid 8. For a detailed discussion on relationship between profiling and automated decision-making, see Isak Mendoza and Lee A. Bygrave, 'The Right Not to be Subject to Automated Decisions Based on Profiling' in Tatiana-Eleni Synodinou and others (eds), *EU Internet Law: Regulation and Enforcement* (2017) 90-91.
[240] Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (n 212) 8-9.
[241] ibid 19-20.

considering the lack of compatibility between the main purpose of the contract and the advertising purpose. The second exception allows EU or national law to form a basis on which such decisions can be made solely by automated means, provided that suitable safeguards are in place. However, there is no sign that the lawmakers — at least at EU level — will give the special green light to the OBA sector, and thus, this should not be the hope of the industry. The least unpractical choice thus seems to be consent. It should be noted that, in this Article, what is required is *explicit* consent. In other words, there is no room for the argument of 'implied consent' — an affirmative act by the user such as hitting a button or manually ticking a box will be required.

One particular point that might give rise to uncertainty in this respect is the potential argument that the current OBA profiling practices should not count as 'decisions' as they are not material enough to have significant impact on the user. Article 22 regulates only automated processing (including profiling) that 'produces legal effects concerning him or her or similarly significantly affects him or her.' This qualifier serves to exclude those profiling practices whose consequences are too trivial to affect the data subject.[242] The question then is whether the impact imposed by OBA on Internet users is comparable to 'legal effects'? If the business model involves price discrimination or denial of service provision, this would obviously have an impact equivalent to 'legal effects'. However, as examined in Chapter 3, most of the negative effects of OBA are intangible, chronic and sometimes collective, having very few immediate implications for individual Internet users. It is therefore at least arguable that Article 22 does not apply to general cases of OBA. The Working Party is of the view that '[i]n many typical cases the

[242] See Bart van der Sloot, 'Decisional Privacy 2.0: The Procedural Requirements Implicit in Article 8 ECHR and Its Potential Impact on Profiling' (2017) 7(3) International Data Privacy Law 190, 201.

decision to present targeted advertising based on profiling will not have a similarly significant effect on individuals'.[243] However, the possibility of 'significant effect' is not completely excluded by the Working Party, who argues that, depending on the 'particular characteristics' of the actual practices, an impact as significant as legal effects might nevertheless exist.[244] While an example of the effects of gambling ads on individuals in financial difficulties is given in the guidelines, the Working Party has not taken a further step to also address the long-term risks as highlighted in Chapter 3. Recital 71 instructs data controllers to 'prevent[], inter alia, discriminatory effects on natural persons on the basis of [sensitive categories]'. It is true that in general, OBA operators would refrain from collecting sensitive data, but certain types of profiling may lead to implicit discriminatory effects based on users' sensitive (or non-sensitive) status. If read in such a broader sense, it would be reasonable to believe that OBA practices fall under the scope of Article 22, although this remains open to discussion.

It should also be reiterated that, as underlined by the Working Party, even when profiling does not constitute 'decision-making', it is nonetheless subject to the general provisions of the GDPR.[245] One practical implication of this is that the data subject *always* reserves the right to object to such profiling for purposes of direct marketing,[246] whether significant impacts are involved, and whichever legal basis the profiling relies upon. Once the data subject exercises this right, the data controller should unconditionally cease any further processing for this purpose, and erase the data at the request of the data

[243] Article 29 Data Protection Working Party, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (n 212) 22.
[244] ibid.
[245] ibid 9.
[246] GDPR, art 21(2).

subject.[247] This provision should not be interpreted as replacing the explicit consent requirement with a softer 'opt-in' or 'allowed-unless-objected-to' regime in a direct marketing context. It simply means that even if the profiling practices in question could be seen not as 'automated individual decision-making' (which is questionable) and could rely on the legal ground of 'legitimate interest' instead of 'consent' (which is also questionable), data subjects would still retain their right to object to such processing.

4.7 Summary: Consent and necessity — Still the best partners?

4.7.1 The GDPR: A complex system featuring consent and necessity

It has been two decades since the adoption of the DPD in 1995, during which time the reality of information technologies have advanced so dramatically that our understanding of the essence of data protection issues has been constantly challenged. In the meantime, people's attitudes towards technology, society and law have also experienced remarkable changes. With the inputs from regulatory authorities, judiciary, academia and advocacy groups, data protection law in Europe has developed to its current complexity. The GDPR is expected to incorporate and translate such developments into the framework legislation. As we have seen, the underpinning principles and mechanisms established by the DPD have remained largely intact in the GDPR, with some of them strengthened. In the course of the analysis of how the GDPR's definitions, principles, legal grounds and special rules will apply to and affect the OBA sector, the interactions between all these elements have contributed to a sophisticated picture of the legal efforts to protect personal data. Still, two discernible characteristics emerge out of the convoluted regime.

[247] ibid arts 21(3), 17(1)(c).

First, the role of consent has been significantly reinforced. This should not come as much a surprise, since consent has constantly been at the heart of data protection controversies. Businesses, individuals and regulators all have different perceptions of what counts as valid consent. Such inconsistencies exist even between regulators from one Member State to another. The divergent propositions regarding whether 'implied consent' can be compliant at all may serve as a good example of the problems. As such, the GDPR clarifies what constitutes 'unambiguous' consent, making it almost impossible to argue that consent can be given in a passive manner. Also, the conditions under which consent can be accepted as valid are also set down in greater detail. Such provisions make it mandatory for the data controller to demonstrate compliance, to separate consent from other contractual matters and to provide the possibility to withdraw consent.[248] All these measures amount to a clearer, higher legal bar for the OBA industry to obtain consent that can produce legal effects.

Consent remains but *one of* the six legal grounds on which processing of personal data can be permitted, but the ubiquity of consent in the GDPR in effect gives it a special status. Consent almost becomes the foremost, unexceptional exception to the restrictions on data processing. With valid consent, the data controller may in general process personal data,[249] including sensitive data,[250] conduct automated individual decision-making[251] and transfer personal data to a third country.[252]

The other noteworthy tendency is that the test of necessity has become even more prevalent. As has been made quite clear above, the ways 'necessity'

[248] ibid art 7.
[249] ibid art 6(1)(a).
[250] ibid art 9(2)(a).
[251] ibid art 22(2)(c).
[252] ibid art 49(1)(a).

features in the GDPR are three-fold: First, it serves as a data protection principle, namely the data minimisation principle; second, it appears as part of a number of legal bases whereby data controllers can lawfully process personal data, such as necessity for performance of contract and necessity for legitimate interests; third, it also operates as a legal test to complement other safeguards in the GDPR, such as consent and purpose limitation. The last point may be less noticeable but in fact makes the biggest difference. For one thing, consent, as analysed above, has explicitly incorporated the necessity element in the assessment of the 'freely-given' factor. For another, the necessity test is also mirrored in the principle of purpose limitation, albeit in a subtler fashion. When evaluating the compatibility of the further purpose with the original one, one essential aspect concerns the relationship between these purposes. The GDPR itself does not require the secondary purpose to be *necessary* for the primary one, but it is reasonable to infer that this would make a strong case for the satisfaction of compatibility. In fact, save that it is expressly provided by law or safeguarded by consent, it is hard to come up with a compatible purpose that is unnecessary for the original one.

That means, even when OBA operators decide to base their processing activities on consent, they would still need to pass two necessity tests: first, whether the data processing envisaged is necessary for its purpose (necessity as a principle, data minimisation); and second, whether the processing is necessary for the performance of contract, if the contract is conditional on the consent (necessity as a legal test). Here, it is important to keep in mind that under EU data protection law, processing of personal data is prohibited as a matter of principle, and only permissible when certain criteria are met. Accordingly, the failure in passing either of the tests would lead to the invalidity of the processing entirely. Such criteria include a set of legal

principles, a set of legal bases, and a set of special arrangements, most of which are now more or less informed by the necessity test under the GDPR.

4.7.2 OBA: Gloomy legal future unless radical changes are made

Now it becomes even more evident how the components of data protection law impact one another. The GDPR, supplemented with other sector-specific legislations, operates in a way that entails an overall understanding from the most abstract principles to the most specific rules, among which certain critical ideas have built strong connections. The EU data protection regime is bolstered by the GDPR's higher consistency and comprehensiveness, and that is probably not good news for the OBA industry. To sum up the legal evaluation conducted in this chapter, if the business practices of the most common operational model of an ad network provider may arguably be considered on the edge of compliance under the current legal framework — the DPD — then there is little chance, at least in theory, that they may survive the enhanced standards under the GDPR.

A number of conclusions can be drawn from the previous sections: (a) The data collected and processed in an OBA network will doubtlessly fall within the scope of personal data under the updated version of definition; (b) Consent will be the only practical option when it comes to the legal basis on which data processing for OBA purposes can be based; (c) The principles of lawfulness, fairness and transparency, data minimisation, and purpose limitation must be observed for all the data processing activities; (d) Special rules, such as those regarding profiling, must also be fully enforced. In practice, the analyses above will have three major implications for the OBA industry.

First, the way consent is obtained needs to be brought further in line with the 'unambiguous' requirement. The so-called 'implied consent', as examined above, would be most unlikely to be considered compliant in the light of the

GDPR's new provisions. At the beginning of the data lifecycle, the collection of user data for tracking purpose, especially by means of cookies, must be warranted by the user's 'opt-in' consent. It would probably not suffice to simply display a banner on the webpage with the notification like 'by using our website, you agree to our privacy policy and use of cookies'. Instead, at least two options, yes or no, should be provided to the user for the first visit to one of the participating websites (ad publishers) of an OBA network. Before the user clicks 'yes', no data processing may take place.

Second, consent for data processing for OBA purposes may no longer be requested as a condition for service provision, not even on an 'opt-out' basis. It follows that OBA operators might be allowed to collect behavioural data only on a fully voluntary basis, or where at least alternative options are provided so individuals are genuinely entitled to 'free choice'. That means, if the user decides to say 'no', they should nevertheless be allowed to use the website, at least for those parts that do not necessitate verification of identity, and with no harm to their user experience, such as obtrusive, repetitive requests of consent.

Third, although real-time bidding or other programmatic trading techniques are not illegal themselves, it is most likely that they will no longer qualify as a compatible secondary use of personal data, if that does not form part of the primary purpose for which the data is originally collected. In other words, these categories of processing should be in most cases treated as initial use and justified by one of the legal bases. The nature and measures of such activities must be clearly stated in the consent request.

Of course, the severity of non-compliance will still depend on how the law is going to be enforced, and almost certainly, the OBA sector will argue for a much weaker approach to these safeguards. However, with the regulators'

supportive interpretations cited throughout this chapter, the purely doctrinal analysis here will almost inevitably lead to the conclusion that the continuation of current practices would be found in breach of the GDPR. This results from all the new measures introduced into the GDPR, which in effect have strengthened the existing regime under the DPD, and give rise to a 'consent + necessity 2.0' model. Within this model, the criteria of valid consent have been enhanced, while the necessity requirement has become more ubiquitous, as a fundamental principle, as legal bases, and as a benchmark for other safeguards.

The implementation of all these new requirements would have little technical obstacle, as will be further shown in Chapter 6. Before considering these practical issues for the perspective of data controllers, it would make more sense to evaluate, from a fundamental right perspective, the extent to which the GDPR can safeguard the freedoms and rights of individuals, as well as the collective well-being of the society. In this Chapter, the focus was mainly on how GDPR *will* regulate OBA, not how it *should*. Therefore, the next chapter will turn to a less doctrinal, but more critical approach.

Chapter 5 Competence of Data Protection Law in the Age of Big Data: A Critical Assessment of EU Data Protection Law in the Context of Online Behavioural Advertising

5.1 Reimagining data subjects

Since the design of any regulatory framework depends largely on how the policymakers assume all the affected parties would behave within certain constraints, it is essential that the enacting of any legislation be based on reasonable assumptions of policy objectives and behavioural patterns. For a very long time during the development of data protection law, there have been a number of such assumptions. For instance, it is assumed that individuals are those who deserve special protection when it comes to automatic processing of information about them.[1] In the same vein, it is also assumed that, provided with sufficient information and genuine choice, individuals would make the best decisions for themselves.[2] As a result, the concept of 'data subject' has been put forward as a legal creation, surrounding which an organic set of rights, duties, mechanisms and remedies have taken shape. The European data protection model was developed — and believed to be effective — at a time when data processing was low-scale, scattered, traceable, costly and time-consuming.

A lot has changed, however, as demonstrated in the preceding chapters. The emergence of big data technologies means that the present-day practices

[1] Council of Europe, Committee of Ministers, Resolution (73)22 on the Protection of the Privacy of Individuals Vis-à-vis Electronic Data Banks in the Private Sector.
[2] Lorrie Faith Cranor, 'Necessary But Not Sufficient: Standardized Mechanisms for Privacy Notice and Choice' (2012) 10 Journal on Telecommunications and High Technology Law 273.

of data processing have become more powerful and more omnipresent. The use of personal data concerns not only *some* aspects of life anymore; it affects nearly *all* aspects of life. It has also presented new challenges to our perceptions of what 'data subject' means — not merely in the sense that these technologies have extended the frontiers of 'identifiable person', but more importantly, in the sense that they have called into question the values that data protection law pledges to protect with the legal identity of 'data subject'. While the GDPR aims to protect 'fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data', how this should be translated into the specific regulatory measures seems to be based on what is at best an abstract idea without conceptual reflection and at worst a hollow repetition. As highlighted in Chapter 3, the risks arising from the use of personal data in the age of big data today are multi-level and multi-faceted, requiring reflections on the policy premises as well as the regulatory approaches. Most importantly, when designing an appropriate framework, it is essential to conceive of data subjects not as isolated individuals, but rather actors with various motives and interests in a highly complex network, interacting with each other, and also other categories of actors.

As will be seen in Section 5.2, data subjects are first and foremost autonomous agents achieving self-actualisation through those mundane decisions that are surrounded by the delicate balance of constraints and choices. To create and maintain the conditions for the fulfilment of one's autonomy, however, it is critical to first revisit the essence of autonomy, in particular in the context of intensive use of personal data today. Besides, data subjects are also engaged in different types of social relationships, which can be of economic or political nature. Considering data subjects largely as consumers, Section 5.3 will recapture the dilemma of choice as highlighted in

previous chapters, with a fresh, critical approach to the implications for policy-making. This will be further explored in a political context in Section 5.4, where data subjects are seen as an organic part of a democratic community. Such a capacity carries a distinct kind of interest that is more collective than individualistic, presumably involving strategies against what one might call the 'tragedy of the commons'.[3]

These inquiries offer a sound analytical foundation for the assessment of the extent to which a given legal framework may or may not address the big issues of big data. Given the diverse nature of the interests and relationships examined throughout this chapter, the analysis of the framework concerned — the GDPR — will be conducted separately in each section. The investigation will draw heavily from the findings in Chapter 3, where the potential risks of big data have been theorised, as well as Chapter 4, where the 'consent + necessity 2.0' model has been extracted from the main body of the GDPR. As the limitations of the GDPR reveal themselves in a range of different contexts, the theoretical, overarching flaws of the current regime will emerge and be summarised at the end of this chapter.

5.2 Retaining autonomy through data protection

5.2.1 Autonomy: Why it matters to self-determination, liberty and equality

In Chapter 3, a range of potential risks arising from the use of personal data have been identified. Some values highlighted there are more individualistic than others, in that they have more concerns about the development of one's identity of 'self', among which are informational self-determination and human dignity. Also, human dignity has been further clarified and

[3] For an introduction to the concept, see David Feeny and others, 'The Tragedy of the Commons: Twenty-two Years Later' (1990) 18(1) Human Ecology 1.

materialised into two essential points: liberty and equality. In this section, an argument will be first advanced that autonomy, from a fundament rights point of view, underpins self-determination, liberty and equality. It will be shown why, in order to safeguard these fundamental values, a data protection framework needs to embrace a critically re-examined notion of autonomy. After that, a more in-depth enquiry into the genuine essence of autonomy in the context of personal data will be conducted, followed by an assessment of whether and to what extent the regime of the GDPR may sufficiently protect the autonomy of data subjects — who are largely considered independent actors in this context.

(a) Autonomy as the ultimate objective of self-determination

The relationship between individualistic autonomy and informational self-determination deserves some further explanation here. Essentially, somebody being autonomous in a society means he or she possesses what it takes to make major decisions in life.[4] As such, minors are commonly not considered fully autonomous,[5] as they are assumed to lack the intellectual maturity to make informed decisions. Slaves are unlikely to be autonomous, as they do not enjoy the necessary freedom to act as they wish.[6] Those who have no disposable property at all can hardly claim to have autonomy as they do not have the minimum instrument to realise their even most limited choice.[7] Among the crucial elements that make a person autonomous, the ability to control their own personal information is becoming more important than ever in a contemporary society. Reminded of what the German Constitutional Court

[4] Joseph Raz, *The Morality of Freedom* (Oxford University Press 1986) 369.
[5] See Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge University Press 1988) 9.
[6] See ibid 129.
[7] See N.E. Simmonds, 'Property, Autonomy and Welfare' (1981) 67(1) ARSP 61, 66.

stated in the case confirming the right to informational self-determination, this point would be even more straightforward.[8] If at that time, the unfettered collection of demographic information about citizens in a national census could be held as potential interference with individual freedom, today's ubiquitous tracking practices on the Internet should be even more alarming.

In the past, people needed privacy, because they would otherwise be subject to constant observation and judgment, which would in turn largely limit their choice of what they can do in a care-free manner. To such an extent, those without privacy live without true autonomy. While these effects all remain significant in today's big data reality, what brings the seriousness to a different level is the fact that, without effective control over personal data, people would not even be aware that decision-making in life might have been compromised. Individuals would in effect be subject to the manipulation of those who get hold of their data if their personal data has been utilised without their oversight. OBA serves as a case in point here. As shown in previous chapters, the marketing industry indeed has the technological power to analyse user behaviour with rather high accuracy and then personalise the content so as to better target their users. In a world without regulation over the use of personal data, individuals would be vulnerable to such influence, and thus, would not be able to make their own decisions in a truly autonomous way.

Informational self-determination thus forms an indispensable part of autonomy, in particular with regard to the activities involving the use of data, which is making up an increasing part of contemporary life. In other words, in the data society that we are currently experiencing, it is impossible to

[8] BVerfGE 65, 1 [1965].

achieve autonomy without effective informational self-determination.[9] However, it should be stressed that there is an important difference between informational self-determination and autonomy,[10] which leads to the more complex issues about the essence of autonomy. Narrowly defined within the 'control' paradigm, informational self-determination mainly concerns one's control over *personal information*,[11] whereas autonomy has more to do with the choice of one's principal course of *life*. Essentially, determining what happens to one's own data is but *one of* the approaches to autonomy. This also echoes a finding in Chapter 3 that informational self-determination has a strong instrumentalist characteristic. Accordingly, it is imperative to keep in mind that informational self-determination serves autonomy, not the other way around. While full control over one's personal data is crucial in many scenarios of life, this is in theory not always the case. As will be shown later in this chapter, unchecked individual control of personal data may sometimes run counter to a broadly understood notion of personal autonomy.[12] For the purpose of this section, however, it would suffice to take note of the crucial link between informational self-determination and autonomy.

---

[9] Eoin Carolan and Alessandro Spina, 'Behavioural Sciences and EU Data Protection Law: Challenges and Opportunities' in Alberto Alemanno and Anne-Lise Sibony (eds), *Nudge and the Law: A European Perspective* (Hart Publishing 2015) 165-166.

[10] See Woodrow Hartzog, *Privacy's Blueprint: The Battle to Control the Design of New Technologies* (Harvard University Press 2018) 118-119.

[11] See Claudia Quelle, 'Not just User Control in the General Data Protection Regulation: On the Problems with Choice and Paternalism, and on the Point of Data Protection' in Anja Lehmann and others (eds), *Privacy and Identity Management: Facing up to Next Steps* (Springer 2016) 144; Sophie C. Boerman, Sanne Kruikemeier and Frederik J. Zuiderveen Borgesius, 'Online Behavioral Advertising: A Literature Review and Research Agenda' (2017) 46(3) Journal of Advertising 363, 374.

[12] Christophe Lazaro and Daniel Le Métayer, 'Control over Personal Data: True Remedy or Fairy Tale?' (2015) 12(1) SCRIPT-ed 3.

(b) Autonomy as the precondition of liberty and equality

A key part of the discussion in Chapter 3 concerns human dignity, which ends up in two strands of enquiries: liberty and equality. By way of recapitulation, it was concluded that the ubiquitous collection of personal data, as demonstrated in the case of OBA, might pose a threat to one's liberty in that the data subject is likely to become self-conscious about being monitored or discriminated, and thus refrain from certain behaviour. In a similar vein, the principle of equality that most progressive societies value would also be eroded because of unfair treatment based on the use of personal data. Such unfairness may on the one hand stem from incorrect information or over-generalised categorisation, and may on the other stem from accurate but uncontested, opaque differentiation.[13] All these potential adverse effects might present serious risks to the way data subjects act as autonomous agents.

When personal data is now being utilised extensively, a helpful approach to maintaining human dignity would need to address what liberty and equality mean in the context of big data. To further explore what the law can and should do with human dignity, it would be helpful to draw a conceptual connection with the idea of autonomy. It is not hard to explain why autonomy lies at the heart of both liberty and equality. As Rawls interprets a Kantian concept of autonomy, 'a person is acting autonomously when the principles of his action are chosen by him as the most adequate possible expression of his nature as a *free* and *equal* rational being.'[14] The values of liberty and equality

[13] For a thorough theory of both kinds of risks, see Jiahong Chen, 'The Dangers of Accuracy: Exploring the Other Side of the Data Quality Principle' (2018) 4(1) European Data Protection Law Review 36.

[14] John Rawls, *A Theory of Justice* (revised edn, Harvard University Press 2009) 222 (emphasis added).

are inherently embedded in the very notion of autonomy.[15] One who acts under the unjustified external influence, whether unfavourable consequences or unfair biases, cannot be said to be acting with autonomy.

(c) The realisation of autonomy

There is a remarkably rich body of literature addressing the very idea of autonomy. Benn, for example, considers an autonomous man as someone 'whose life has a consistency that derives from a coherent set of beliefs, values, and principles, by which his actions are governed.'[16] For him, choice and rational criticism are the necessary conditions for autonomy,[17] although more emphasis has been placed on the latter part, with the statement that '[t]o be a chooser is not enough for autonomy'.[18]

Raz, with a somewhat different approach, captures the essence of personal autonomy as 'the vision of people controlling, to some degree, their own destiny, fashioning it through successive decisions throughout their lives.'[19] For an agent to be autonomous, three conditions must be met: minimum mental capacities, adequacy of options and independence from coercion.[20] Clearly, these conditions imply something more objective and external than those put forward by Benn. What matters to the achievement of autonomy, for Raz, is the 'environment' in which autonomous life may flourish.[21] Accordingly, a desirable model of political freedom should be one that 'protects people pursuing different styles of life from the intolerance […], and

---

[15] For a different theory about the relationship between liberty (or freedom) and autonomy – one that considers liberty as an important part but not all of autonomy – see Dworkin (n 5) 12-20.

[16] S. I. Benn, 'Freedom, Autonomy and the Concept of a Person' (1975) 76 Proceedings of the Aristotelian Society 109, 124.

[17] ibid 127.

[18] ibid 123.

[19] Raz (n 4) 369.

[20] ibid 369-378.

[21] ibid 391.

it calls for the provision of the conditions of autonomy without which autonomous life is impossible.'[22]

Between these two approaches sits Dworkin's theory of autonomy, which has a strong internal focus on individuals' ability to reflect on decisions in life, and yet opens up an external avenue for critical assessment of political institutions. He considers autonomy as an intermediary between one's specific ('first-order') preferences and the general, principled ('higher-order') values possessed by the person. For him, 'autonomy is conceived of as a second-order capacity of persons to reflect critically upon their first-order preferences, desires, wishes, and so forth and the capacity to accept or attempt to change these in light of higher-order preferences and values. By exercising such a capacity, persons define their nature, give meaning and coherence to their lives, and take responsibility for the kind of person they are.'[23] Within the particular debate over freedom of expression, Dworkin spells out the practical implications of such a capacity:

> 'a state may be required to recognize political autonomy of its citizens. That is, it may not restrict the liberty of individuals unless it can justify such restrictions with arguments that the person himself can (given certain minimal rationality) see as correct.'[24]

Apart from such classical theories about autonomy, there are also alternative theoretical frameworks applicable to particular contexts, including those concerning the use of personal data.[25] Interesting as this could be, it is not the task of this chapter to delve into all these philosophical theories. With in mind the major objective of this chapter — a critical assessment of the GDPR — Raz's theory is chosen as the analytical framework, for a number of reasons.

---

[22] ibid 425.
[23] Dworkin (n 5) 20.
[24] ibid 40.
[25] See, for example, Paul Bernal, *Internet Privacy Rights: Rights to Protect Autonomy* (Cambridge University Press 2014); Quelle (n 11).

The theory has a generic nature that can by and large accommodate the common ground shared by alternative theories, which, for example, usually see autonomy as a matter of way of life instead of particular decisions. Also, as a political theory of law, Raz's account — in particular his later works, as analysed below, provides a particularly useful approach to the evaluation of a legal instrument. Its emphasis on the role and limitations of the authority of law in promoting autonomy fits nicely into the ongoing debate over the regulatory model of data protection law.

An important part of Raz's work on theories of law concerns the seemingly contradictory normative requirements between human reason and authority. His question, in short, is that, if the very nature of the authority of law demands the obedience by its subjects without questioning the rationale, then how can this be compatible with the autonomy of mankind, which relies on the reasoning of oneself?[26]

His answer is complex yet straightforward. In everyday life, it is not uncommon that we give up the right to make the final decision or limit our future choice by, for example, committing ourselves to a contract, setting a speed limit and so on.[27] The point is that there are other ways to achieve the primary values of the ability to act by our own judgement in conformity with reasons.[28] For Raz, subjecting oneself to an authority is 'not a denial of people's capacity for rational action, but simply one device', if the authority enables its subjects to better conform to reason.[29]

There are of course cases where such commitments cannot be considered a device for rationality. A promise to be a slave, for instance, would not count

[26] Joseph Raz, *Between Authority and Interpretation: On the Theory of Law and Practical Reason* (Oxford University Press 2009) 135.
[27] ibid 140.
[28] ibid.
[29] ibid.

as an act of autonomy based on rationality.[30] An analogy can be made in a data protection context: One who makes blanket promise to allow the use of all their data for all purposes would not be considered to have exercised their autonomy. To qualify as an enabling instrument, an act 'must belong to a class of actions such that it enhances people's control over their life'.[31] Clearly, the essential element of a Razian version of autonomy consists in the ability to, through certain mechanisms, get hold of greater powers on future life. Such greater powers can be gained or protected by, for example, putting oneself in a position to be bound by certain restrictions in the future, or by alienating part of one's freedoms to a collective enterprise in quest of greater potentials that cannot be achieved individually. The former can be exemplified by contractual arrangements, including promise and consent, and the latter by social or institutional schemes, such as legal regulation. These mechanisms are also readily available as part of a typical data protection legal framework.

In short, the key idea of Raz's autonomy-based theory of law is that, as a practical authority, the justification of law lies in its potential to facilitate people in achieving greater control over life, which is in line with, not contrary to, the essence of autonomy. Of course, such a general idea needs to be elaborated in more detail and put to the test with instances of legal systems. The 'consent + necessity 2.0' model of the GDPR will be evaluated along with the refined notion of autonomy in the following sections.

5.2.2 Consent — What does control mean today?

(a) The merits of consent

If the degree of possible individual control over personal data forms an essential part of autonomy promoted by law, an assessment of a legal

[30] ibid 136.
[31] ibid.

framework would need to investigate the extent to which data subjects are entitled to determine how their data is processed. Consent as a mechanism to enable such a possibility has its own merits. Indeed, consent is closely related to the concept of autonomy. A theory of consent, known as 'autonomous authorisation model', elaborates the connection by stating that '[o]ur right as rational human beings to choose our course of action freely is reflected in our ability to consent.'[32] Ideally, as a mechanism to authorise the performance of an action, consent gives the involved party an opportunity to contemplate the trade-offs and decide whether to accept or decline the offer. A properly designed consent-based system would represent an informed expression of one's will, which involves 'agency, volition, control deliberateness and making something happen.'[33] Without all these conditions, the fact that somebody says 'I agree' might happen simply as a result of coercion. This is why meaningful consent would essentially entail the act of a positive confirmation, and also why the idea of 'implied consent' is intrinsically incompatible with the spirit of 'unambiguous consent'. As pointed out in the previous chapter, the provisions in the GDPR concerning consent have been updated so as to specify what amounts to sufficient, valid consent. As a consequence, the argument of 'implied' or 'opt-out' consent is unlikely to be compatible with the GDPR.

Therefore, the new changes regarding the concept of consent in the GDPR may probably help bring in new practices of obtaining better-informed consent that is in line with the purpose of protecting personal data. Having said that, such improvements, though helpful, are not necessarily sufficient in

[32] Bart W. Schermer, Bart Custers and Simone van de Hof, 'The Crisis of Consent: How Stronger Legal Protection May Lead to Weaker Consent in Data Protection' (2014) 16 Ethics and Information Technology 171, 174.
[33] Andrew McStay, 'I Consent: An Analysis of the Cookie Directive and Its Implications for UK Behavioral Advertising' (2012) 15(4) New Media & Society 596, 600.

protecting individual autonomy. The effectiveness of consent is at least disputable from various angles, as will be shown in the remaining sections of this chapter. Here, even if we assume that the legal framework can in effect ensure all the conditions of consent are met — namely 'freely given, specific, informed and unambiguous'[34] — certain issues would nevertheless remain: the rigidity and flatness of consent.

(b) Rigidity of consent

To understand what rigidity means and why it matters in this context, it is useful to refer to some theories of contract law. Freedom of contract, as a general principle, allows contracting parties to freely decide what the rights and duties are under a valid contract.[35] The implications of this principle are twofold. First, government should in principle refrain from imposing restrictions on the content of contracts; second, contrary to the principle of *numerus clausus* in property law, which prohibits the parties of a transaction to create new types of property rights other than those prescribed by law,[36] the content of the clauses is not limited to the typical forms of contracts. Of course, contemporary theories and practices have evolved significantly and it is commonplace today for law to impose restrictions on, for instance, contracts of adhesion, and to provide certain standardised clauses for typical contracts in the case of ambiguity.

---

[34] 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] OJ L119/1 ('GDPR'), art 4(11).

[35] Richard A. Epstein, 'Contracts Small and Contract Large: Contract Law through the Lens of Laissez-Faire' in F. H. Buckley (ed), *The Fall and Rise of Freedom of Contract* (Duke University Press 1999) 28.

[36] Thomas W. Merrill and Henry E. Smith, 'Optimal Standardization in the Law of Property: The *Numerus Clausus* Principle' (2000) 110(1) The Yale Law Journal 1.

With regard to data processing, consent is no different from a contract in other areas — it creates rights and obligations in the form of mutual consensus. However, in practice, particularly in an online setting, consent almost always turns up as a contract of adhesion — the way Internet users give consent to processing of their online data is completely the same as concluding a take-it-or-leave-it contract. It is not (only) to say that consent is not freely given, which is a separate issue to deal with. The problem being discussed here is rather the technological architecture of consent: inflexible and indivisible. Taking OBA as an example, it is quite conceivable that some Internet users would have no issue allowing first-party tracking but would rather stay away from third-party tracking. Even for those who do not completely reject the idea of third-party tracking, they might want to keep apart two categories of profiles. Of course, all these different uses of personal data can in theory be seen as separate operations on the same set of personal data, and thus each requires a separate confirmation by the data subject. However, these options are far from available in reality, and for an average user, consent works like a once-and-for-all yes-or-no answer, despite the possibility of withdrawing consent. Consent essentially only indicates an overall acceptance or rejection of a package, and thus cannot fully realise the wide range of possibilities.

The lack of opportunities for data subjects to precisely calibrate the use of their data was a lesser problem in the 1990s. Data technologies then were not sophisticated enough to exert an all-dimensional influence on people's life. In today's age of big data, however, as the impact of use of personal data has become omnipresent, the need for more subtle ways to exercise individual control over personal data has become more pressing. Autonomy is a fluid concept and depends on the socio-technical reality of the society. Selective disclosure of information to others has always been common in human

history.[37] A piece of yes-or-no consent given to individual websites might have sufficed to protect autonomy two decades ago, given the state-of-the-art then. In the face of big data, however, this can no longer satisfy individuals' need to exercise their autonomy in a more individualised manner. The possibility to individually 'customise' how personal data is collected, analysed and shared now should now be considered an indispensable part of informational self-determination that deserves further protection.

Of course, this would not solve the problem overnight. It is conceivable that when such measures are first introduced, many users would not be aware of their availability. Even if they are aware, chances are that they do not have the necessary skills, time and motivation to make the most of such functionalities. Even worse, this could give data controllers the excuse to introduce more intrusive default settings, as they may argue the users have been given sufficient choice. However, this can be at least partly prevented by setting down a robust set of rules for default settings (the criteria of which will be further discussed later in this chapter). Also, it is true that the adoption of such measures by Internet users might take time, but this does not mean that it should never happen. The popularity and ease of the use of such measures, as will be explored in Chapter 6, can be facilitated by certain technological, market and legal solutions.

In fact, some service providers have already deployed what they call a 'privacy dashboard' to allow users to indicate precisely how they want their data to be used in different settings.[38] Powered by big data, providing such functionalities has become technically more feasible, and financially less costly.

---

[37] Jan Van Dijk, *The Network Society* (3rd edn, SAGE 2012) 217.

[38] Elyse Betters, 'New Windows 10 Privacy Dashboard Gives You More Control over Data' *Pocket-Init* (11 January 2017) <http://www.pocket-lint.com/news/139971-new-windows-10-privacy-dashboard-gives-you-more-control-over-data> accessed 25 July 2017.

Again, industrial initiatives like these are welcome but much more can be done. Presumably, a robust, holistic approach to lower costs for data subjects to effectively benefit from more nuanced control would entail not just legal solutions, but also those of technology and market. While it is true that the lawmakers of the GDPR seems to some extent aware of the potential of the non-legal approaches and have laid down such provisions as the 'data protection by design' requirement, it is entirely unclear how the rigidity of consent could be addressed.[39]

(c) Flatness of consent

Another limitation of consent, as it currently stands in the GDPR, is its 'flat' nature, which refers to its inability to signal the essentiality or triviality of the matters concerned. It is well proved in a good number of empirical studies that Internet users rarely read privacy policies provided by online service providers,[40] and give consent to these services regardless of the content.[41] There are a number of common accounts for the failure of consent in this regard. First, Internet users have been exposed to too much an amount of consent requests, causing what is called 'consent fatigue'.[42] Second, privacy policies are usually overloaded with excessive information and details, which would create tremendous burdens on the users had they read all these

[39] See Bert-Jaap Koops and Ronald Leenes, 'Privacy Regulation Cannot Be Harcoded. A Critical Comment on the "Privacy by Design" Provision in Data-Protection Law' (2014) 28(2) International Review of Law, Computers & Technology 159, 166.
[40] See, for instance, Carlos Jensen, Colin Potts and Christian Jensen, 'Privacy Practices of Internet Users: Self-reports versus Observed Behavior' (2005) 63 International Journal of Human-Computer Studies 203; Rainer Böhme and Stefan Köpsell, 'Trained to Accept?: A Field Experiment on Consent Dialogs' (Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, 10-15 April 2010); Idris Adjerid and others, 'Sleights of Privacy: Framing, Disclosures, and the Limits of Transparency' (Proceedings of the Ninth Symposium on Usable Privacy and Security Newcastle, 24-26 July 2013).
[41] Paul M. Schwartz, 'Internet Privacy and the State' (2000) 32 Connecticut Law Review 815, 824.
[42] Schermer, Custers and Hof (n 32) 176.

notices.[43] Third, there is usually a lack of genuine choice for individual users and thus they are virtually forced to agree to the use of their data.[44]

This section will focus on the first two issues, namely overexposure to consent requests, and overcomplexity of consent specifications, while the next section will conduct a more detailed enquiry into the third. Neither of the two former scenarios is limited only to an online setting. We engage in all sorts of contractual agreements as part of daily routine. While it is unquestionably advisable to read terms carefully before signing a deed to transfer a real estate, very few people actually read the fine print on the back of a shipping receipt when dropping a parcel for delivery. Presumably, people are much less willing to read the lengthy legal terms in the latter cases for two main reasons. First, the consequences of ignorance of the terms and conditions of small transactions are too minimal compared to the time and efforts that reading them all would take. Second, most people trust that their basic rights as consumers are protected by statutory measures and cannot be excluded even by mutual consensus. Depending on the seriousness of a transaction, taking actions with different levels of caution is entirely reasonable. Such strategies can hardly be deemed as a compromise of one's autonomy. To the contrary, within the constraints of time and resources, such behaviour can be an example of exercising autonomous rationality. This is because the actors in these scenarios may have in effect freed themselves from the time-consuming task of reading through everything, and thus, in Raz's terms, have enhanced the control over their life.

---

[43] Aleecia M. McDonald and Lorrie Faith Cranor, 'The Cost of Reading Privacy Policies' (2008) 4(3) I/S: A Journal of Law and Policy 543.

[44] Federal Trade Commission, 'Protecting Consumer Privacy in an Era of Rapid Change: A Proposed Framework for Businesses and Policymakers (Preliminary FTC Staff Report)' (2010), 52.

The implications of such a philosophical reflection of consent and autonomy are two-fold for data protection. For one thing, maintaining autonomy does not always involve requiring people to make and carry out all the decisions themselves. Instead, autonomy can also be achieved by delegating decision-making to other agents. Certainly, that would need to be based on some kind of reasonable guarantee; otherwise, it would just be blind trust. For another, even under circumstances where the decision is to be made by the person concerned, the burden of precaution on the decision-maker should be proportionate to the potential consequences of the decision. The more important the decision is, the greater attention the person involved should pay. In the context of consenting to data processing online, the same principle holds true: A mixture of varying levels of priorities concerning data processing means that there should be a hybrid of different approaches to consent.

Unlike handling most decisions offline, however, there are several problems with consent in an online environment. Average Internet users lack the ability to tell the important consent requests from the unimportant ones.[45] Worse still, even if they were able to make such a distinction, alternative mechanisms of giving consent would unlikely be available. That is why some researchers suggest that 'implied consent' be introduced as a new legal basis for low-risk transactions.[46] A similar approach is also recommended by the US FTC, proposing that 'affirmative express consent' should be required only in such cases of material change to privacy policies or use of sensitive data.[47] Of course, it would be highly challenging for the public (or their delegates) to

[45] Daniel J. Solove, 'Privacy Self-Management and the Consent Dilemma' (2013) 126 Harvard Law Review 1880, 1891-1893.
[46] Schermer, Custers and Hof (n 32) 181.
[47] Federal Trade Commission, 'Self-Regulatory Principles for Online Behavioral Advertising: Tracking, Targeting and Technology' (2009) FTC Staff Report, 41-44.

debate and decide what constitutes 'low-risk' activities. From a legislative point of view, this would be even more complicated when it comes to *who* to decide. Yet, there is a clear demand in today's complex digital landscape to expand consent from one single measure to a spectrum of measures. Unfortunately, the GDPR seems to have only tried to strengthen a strictest sense of 'unambiguous/explicit consent' and have not considered the possibilities of the softer alternatives.

Allowing a flexible scheme to tackle the flatness of consent would, again, certainly face with some (fair) scepticisms and even criticisms. The biggest practical issue, of course, would be how to define 'low-risk' and 'high-risk' activities, and how to avoid potential abuse of such flexibility. As mentioned above, what should be considered high- or low-risk, and who should decide, would almost inevitably be subject to constant debates. In the next chapter, one potential solution will be proposed, which outlines how legislative, judicial and supervisory bodies may co-create an environment where such norms can be developed and market actors can be motivated to adhere to these norms. Here, it would suffice to note that there should be room for the members of the society to discuss and decide how a range of regulatory possibilities can be utilised. In such a course, baseline safeguards should always be in place before consensus can be reached. In particular, it is essential to maintain trust by making sure the basic rights of individuals cannot be waived even by consent. Reaching social consensus would take time, but it pays to create an adaptive system to replace the flat model of consent, as the latter has been proved unhelpful in achieving the genuine autonomy.

5.2.3 Necessity — How mandatory standards may help free choice

If it is true that the authority of law is not necessarily contrary to what autonomy entails, and even that sometimes the achievement of autonomy

depends on the mandatory requirements of law, then the next question would be where the line should lie. Labour law, for example, is well-known for precluding the consensual discretion of the employers and employees on certain matters. These stipulations are often controversial as to whether they are in effect enabling or restricting the freedom of the employees. In a similar vein, data protection law sometimes imposes additional restrictions on certain categories of data processing irrespective of the existence of valid consent. The data protection framework in the EU, whether the DPD or the GDPR, has adopted a 'double safety valve' architecture: data processing must be carried out on at least *one of* the legal bases, and at the same time, in conformity with *all of* the legal principles. That means some processing on personal data, even warranted by fully valid consent, may nevertheless be held unlawful if it is found to have breached one of the principles. To such an extent, these principles altogether form a set of minimal requirements that cannot simply be waived by consent.[48]

To further enable Internet users to effectively maintain their autonomous power, the minimum requirements set out by law should avoid two extremes: not to be too high for individuals to decide for themselves, and not to be too low for real choice to be meaningful. When it comes to the online environment, the scope should be such that the defects of consent can be overcome yet the restrictions do not go beyond that. Considering the problems of consent identified in the previous section, the mandatory safeguards should discharge two missions. First, to find out and prohibit those categories of data processing that would deprive individuals of their decision-making power on matters that might gravely endanger their fundamental autonomy. Data processing

[48] For a similar idea that contrast consent and principles as two regulatory paradigms, see Quelle (n 11) 44-46.

with serious negative externalities, as analysed in the next sections, should also be on the blacklist. Second, in terms of other categories that are more or less permissible, to provide individuals with information about the seriousness of each category and implement proper consent mechanism accordingly.

For the first task, the GDPR has put in place a number of strict prohibitions, some of which are reflected in the data protection principles. For example, the principle of purpose limitation sets forth a few restrictions on the purpose for which the processing is carried out. Pursuant to this principle, 'blanket' consent for all purposes would not be allowed as this can hardly be 'specific' and 'explicit'.[49] Also, it is required that the purpose must be 'legitimate'.[50] As such, processing of personal data for illegitimate purposes, such as implementing differential treatments forbidden by anti-discrimination laws, would not be lawful even if the data subject has expressly consented. Article 9, though not laid down formally as a principle, stipulates that processing of sensitive data be prohibited in principle.[51] There are of course other context-specific types of data processing that are undesirable no matter if consent has been obtained. For example, use of personal data for political campaign, as will be discussed in Section 5.4, might be one of those types. It remains to be seen how effectively such safeguards will be enforce, but it seems that the GDPR has already taken note of some of the highest-risk activities and afforded certain safeguards accordingly.

Regarding the second task — flagging up the level of importance to data subjects when they are to decide whether and how to give consent — the GDPR does not seem to live up to its job. The way, and the only way, consent

[49] GDPR, art 5(1)(b).
[50] ibid.
[51] ibid art 9.

can be recognised as valid would be to ensure it is 'freely given, specific, informed and unambiguous'.[52] Of course, the data controller may rely on other legal bases, all of which are characterised by the one-size-fits-all necessity test. The distinction between consent and necessity, as it currently stands in the GDPR, is very clear-cut. Data controllers as well as data subjects have only two options to make a deal: either through fully compliant consent, or through passing the high legal bar of necessity.

This echoes the 'flatness' criticism of consent in the GDPR — necessity is as flat as consent. As summarised in the preceding chapter, the strengthened necessity test underlies the entire GDPR in different forms: as a principle, as legal bases and as additional safeguards. Operating within the data minimisation principle, the necessity test precludes the possibility of data processing that is less 'necessary' but potentially beneficial. In other words, necessity is considered a matter of yes or no under the principle of data minimisation, not one of degree. Such flatness also manifests itself in legal bases other than consent. Whether 'necessary for performance of contract' or 'necessary for legitimate interest', data processing is only allowed when such purposes are, according to the CJEU, 'strictly necessary'.[53] Moreover, when supporting other safeguards, such as the 'freely given' element of consent, the necessity test is likely to invalidate bundled consent for processing unnecessary for the provision of service. There are of course many cases where such a prohibition is crucial, in particular when the data controller exploits its market power to force data subjects to give consent as a condition for access to service. However, there could also be cases where the processing has very low risk and the potential benefits are high, and the data subject truly wishes to

[52] ibid art 4(11).

[53] See ch 4, s 5.C above, citing Joined Cases C-92/09 and C-93/09 *Volker und Markus Schecke and Eifert* [2010] OJ C 13/6 para 77.

accept bundled consent. In these cases, the data subject might lose the opportunity to engage in such a binding arrangement, since the data controller would not offer such a possibility in the first place knowing that such consent would not be enforceable. This way, the strengthened necessity element in the GDPR, as a principle and a benchmark, might end up compromising the autonomy of the data subject by restricting their choice.

It bears repeating that using the necessity test to block unnecessary data processing is *most of the time* a precondition for the realisation of autonomy, but autonomy is also *sometimes* achievable only when certain unnecessary activities are permissible through mutual consensus. The lack of the second possibility under the GDPR will render it counterproductive in protecting autonomy. The 'take-it-or-leave-it' practice adopted by a lot of data controllers are often criticised for undermining individuals' autonomy,[54] but when it comes to what vehicles the law can offer, hardly has this been considered as part of the data protection legal reform.

5.3 Promoting choice through data protection

5.3.1 Choice as an interpersonal interest in commercial contexts

With an overview of the major business models in the OBA sector, Chapter 1 has sketched out a rough picture of how the key players have grown remarkably in size, breadth and impact. One point highlighted in that chapter was that although the ecosystem in this field remains quite diverse, the most powerful giants, with their long-reaching arms, have in fact become increasingly powerful, and consumers are left with a decreasing level of choice. If consumer welfare constitutes an important consideration of data protection

[54] Paul M. Schwartz, 'Privacy and Democracy in Cyberspace' (1999) 52 Vanderbilt Law Review 1609, 1662; Schermer, Custers and Hof (n 32) 176.

law,[55] the evaluation of a legal framework would inevitably have to assess the adequacy in protecting consumer interests. 'Consumer choice' is probably a good starting point here as it is a joint concern shared by consumer protection and data protection laws.[56] From the perspective of consumer protection, consumers benefit from a wider range of choice of goods or services, since intensified competition induces improvements in quality and reduction in price.[57] With few competitors in an area, consumers would conceivably be subject to conditions less favourable than those were they able to switch from one provider to another. As for data protection law, choice matters in that the ability to prevent lock-in to a particular data controller constitutes an essential prerequisite for genuinely 'freely given' consent.[58]

Here, choice represents a category of interest worth promoting through data protection law. It forms part of the bigger picture of individual autonomy since the adequacy of options constitutes a prerequisite for a true sense of autonomy.[59] That said, the analysis in this chapter will take a slightly different approach due to the unique context and interest involved when individuals are engaged in commercial activities and protected as consumers. The context here entails a wider lens beyond the individualistic focus on the construction of one's identity. It takes an alternative perspective through which individual interests are put in the light of interpersonal activities. When commercial

[55] For a discussion about the interactions between data protection, consumer protection and competition law, see European Data Protection Supervisor, 'Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy' (2014) Preliminary Opinion of the European Data Protection Supervisor.
[56] ibid 34-36.
[57] Stephen Weatherill, *EU Consumer Law and Policy* (Elgar 2013) 29.
[58] GDPR, Recital 42.
[59] Raz, *The Morality of Freedom* (n 4) 373.

actors come into play, the values at stake would be better framed vis-à-vis business powers.

According to the findings in Chapter 1, the possibility of real choice in an online setting is being jeopardised at the moment. In fact, the insufficiency of choice in the case of OBA, or even the broader one of the whole Internet, can be understood in three slightly different but interrelated senses: First, there is a lack of alternative services. In a number of the most used online service sectors, such as search engine and video streaming, some service providers are the only dominant ones. Internet users are then, in effect, left with no alternative choice as the most needed resources are likely to be available only on these services. Second, there is a lack of alternative data processing models. While a number of competing business models like 'paid-for' or 'freemium' are gaining popularity in some sectors,[60] 'free' services supported solely by advertising remain the mainstream on the rest of the Internet. In the social media sector, for example, major service providers remain dominantly funded by marketing revenues.[61] For consumers, that means even though they are free

[60] Mark Sweney, 'Online Paid-content Market Poses Threat to Traditional Advertising' *The Guardian* (1 November 2012) <https://www.theguardian.com/media/2012/nov/01/online-paid-contect-rise-8-billion-pounds> accessed 25 July 2017.

[61] Breakdowns of revenues for the top 5 social network platforms on smartphones (as per 2016 Nielsen Social Media Report, in terms of unique users, all included also in the same report's chart on top 10 platforms on PCs): Facebook, Instagram (both owned by Facebook, Inc., advertising: 98.3%, other: 1.7%), Twitter (advertising: 86.3%, other: 13.7%), Pinterest (advertising: 100%), LinkedIn (recruitment service: 62.8%, advertising: 19.4%, subscription: 17.8%, as of the end of 2015). See Nielsen, *2016 Nielsen Social Media Report – Social Studies: A Look at the Social Landscape* (2017) 7 <www.nielsen.com/content/dam/corporate/us/en/reports-downloads/2017-reports/2016-nielsen-social-media-report.pdf> accessed 20 November 2017; Facebook Inc., *Form 10-K* (2018) <http://d18rn0p25nwr6d.cloudfront.net/CIK-0001326801/c826def3-c1dc-47b9-99d9-76c89d6f8e6d.pdf> accessed 8 June 2018; Twitter Inc., *Form 10-K* (2018) <http://files.shareholder.com/downloads/AMDA-2F526X/6318657145x0x972387/4CA553F3-44F1-48CB-8CB5-F03C6FDE95FE/2017_Annual_Report.pdf> accessed 8 June 2018; Kurt Wagner, 'Pinterest Expects to Make More Than $500 Million in Revenue This Year' *Recode* (21 March 2017) <https://www.recode.net/2017/3/21/14991260/pinterest-advertising-revenue-500-million-growth-ipo> accessed 20 November 2017; LinkedIn Corporation, *Form*

to switch to another service, their data would remain to be treated in a largely same manner. Third, there is a lack of alternative data networks. Recalling a diagram in Section 1.3.3 (Figure 1.2) showing how Google has invested on each link of the OBA value chain, it is clear that switching to a different service at the front-end does not necessarily stop data processing within the same network, and thus by the same group of controllers. To be fair, not all of the issues raised here result solely from the use of data and can be fully addressed only by data protection. The intensive use of personal data, however, has indeed exacerbated the problems and to such an extent, a robust data protection framework should acknowledge such challenges and attempt to tackle them as far as possible.

5.3.2 Consent — Simplified choice and complex networks

The way Internet users exercise their choice is, supposedly, to retreat from a particular service, or to switch to a new one. The GDPR provides a mechanism for data subjects to stop their data from being used. As analysed in Chapter 4, data subjects have the right to withdraw their consent at any time.[62] Once the consent is withdrawn, the data controller would not be allowed to process the personal data concerned any longer, and the data subject may also request the erasure of such data, if there is no alternative legal basis for the processing.[63] Also, under the newly-introduced 'right to data portability' provision, data subjects are entitled to have their data transferred directly to another data controller.[64] It is believed that this right would further enable data subjects to realise their choice, as it reduces the switching costs and thus alleviates

10-K (2016) <https://s21.q4cdn.com/738564050/files/doc_financials/annual/2015/LinkedInAnnualReport_2016.PDF> accessed 20 November 2017.

[62] GDPR, art 7(3).

[63] ibid art 17(1)(b).

[64] ibid art 20.

consumer lock-in.[65] These improvements are helpful in facilitating Internet users to retreat or switch from online services, but it remains to be further investigated whether the three 'choice' issues outlined above can be sufficiently addressed.

(a) Lack of alternative services

For the first issue concerning the absence of alternative services in certain sectors, it seems that regulatory instruments within the realm of data protection law has hardly an effective role to play. While the diversity of services is of significant importance, it seems to be more of a competition law issue, falling outside the scope of data protection law. In recent years, the interactions between competition and data protection laws have sparked increasing scholarly and regulatory discussion.[66] It is suggested, for example, that when handling approval of acquisitions and mergers, the regulatory authority should consider not only the firm's market share, but also its data power.[67] In Germany, for instance, the competition authority has found Facebook's excessive collection of personal data an abuse of its market dominance.[68] Also, the breach of commitment regarding the use of personal data may lead to penalties pursuant to competition law. In fact, in May 2017, Facebook was fined €110 million for integrating Facebook and WhatsApp user

---

[65] Orla Lynskey, 'Aligning Data Protection Rights with Competition Law Remedies? The GDPR Right to Data Portability' (forthcoming) European Law Journal.

[66] See, for example, European Data Protection Supervisor (n 55); Wolfgang Kerber, 'Digital Markets, Data, and Privacy: Competition Law, Consumer Law and Data Protection ' (2016) 11 Journal of Intellectual Property Law & Practice 856; Nicolo Zingales, 'Between a Tock and Two Hard Places: WhatsApp at the Crossroad of Competition, Data protection and Consumer law' (2017) 33(4) Computer Law & Security Review 553.

[67] Nathan Newman, 'Search, Antitrust, and the Economics of the Control of User Data' (2014) 31(2) Yale Journal on Regulation 401.

[68] Douglas Busvine, 'Facebook Abused Dominant Position, Says German Watchdog' *Reuters* (19 December 2017) <https://uk.reuters.com/article/us-facebook-germany-dataprotection/facebook-abused-dominant-position-says-german-watchdog-idUKKBN1ED10J> accessed 1 March 2018.

data, contrary to their promise to keep those profiles apart prior to the merger.[69] In this regard, data protection law may of course function as a useful assessment framework as well as compliance guidance to complement competition law. However, any regulatory efforts directly aimed at market structure would after all come down to a matter of competition law, and the role of data protection law would be very limited.

(b) Lack of alternative data processing models

Perhaps data protection law is better positioned to tackle the other two issues. For the second challenge, the lack of alternative data processing models, there is a lot more the legal framework can do. As mentioned above, in some homogeneous market, the only available business model, despite the co-existence of competing services, is keeping their services free of charge but monetising user behavioural data. If the diversity of choice — in terms of how consumers may 'trade' their data — is something desirable or even vital to consumer interests, and if such diversity suffers market failure, the law should arguably encourage, or even impose it. Article 7(4) — which aims to 'unbundle' consent for processing unnecessary for the performance of contract — seems to have some potential on this matter, but does not hold enough promise. As analysed in the previous chapter, this provision has effectively 'injected' the necessity element into the consent requirement. More specifically, if the concerned processing proves unnecessary for the performance of contract, the 'freely given' status could be challenged.[70] The wording 'utmost account' does not say much about the consequences, although Recital 43 provides that such consent is 'presumed not to be freely given'.

---

[69] Jennifer Rankin, 'Facebook Fined £94m for 'Misleading' EU over WhatsApp Takeover ' *The Guardian* (18 May 2017).
[70] GDPR, art 7(4).

As pointed out before, it remains highly uncertain whether the necessity test here connotes technical or financial necessity, or anything else. Technically, OBA is not necessary for most online services, but financially, as many marketers might argue, OBA could be crucial for the operation of such services — although this is at least questionable, as will be shown in the next chapter. One possible interpretation of this provision, in the light of the aim to promote diversity of choice, can be that, where technically possible, the data controller should offer the option of paying (in pecuniary form or otherwise) for not being tracked on that service, as a substitute for allowing use of data for direct marketing purposes based on consent. A potential criticism can be those who can afford paying for tracking-free services would enjoy a higher degree of privacy, while those who cannot would end up with no choice but giving consent. The potential divide of the privacy haves and have-nots is indeed an authentic risk that should be taken seriously in the first place. Accordingly, personal data that is of sensitive nature or data processing that would impose a disproportionate high risk should always be either prohibited or at least unbundled from service provision.

(c) Lack of alternative data network

The last challenge, concerning the behind-the-scenes complex data network, involves the effective execution of withdrawal of consent and similar safeguards. Once the data subject withdraws the consent, the data controller has the duty to erase the data concerned without undue delay,[71] unless there remains another legal ground for the processing or one of the exceptions applies.[72] At the same time, if the data has already been made available to any third parties and the data subject requests the erasure by such third parties as

[71] ibid art 17(1)(b)
[72] ibid arts 17(1)(b), 17(3).

well, the data controller should 'take reasonable steps, including technical measures,' to inform them of the request.[73] More importantly, the data controller should ensure that withdrawing consent should be as easy as giving it.[74] These new stipulations will be helpful, at least in theory, in enabling Internet users to switch from a service belonging to an advertising network to another belonging to a different network. It is noteworthy that the GDPR has in particular underlined the approach of 'technical measures'. Now that big data has made it much easier for data controllers to benefit from the use of data, it should also make it easier for data subjects to exercise their rights. In fact, there have already been initiatives within the online marketing sector to allow users to opt out of OBA in a 'one-stop-shop' manner.[75] It clearly shows that there is no technical obstacle for data controllers to allow data subjects to withdraw consent network-wide in one go. However, if the two services are part of the same OBA network, as the case might well be, switching between them would make little if any difference. That is because by switching to a new service, the user is very likely to have given new consent to carrying on the processing. In the case of Google, for instance, unless a user terminates their account for all Google services (including Google Maps, YouTube, Gmail, etc.), simply switching from a third-party service to another within Google's immense network would hardly reduce the amount of data Google holds about that user.

In order to deal with such a problem, perhaps there should be legal barriers to make the integration of such advertising networks much harder, not just in terms of competition law, but also data protection law. If one recalls the ways

[73] ibid art 17(2).
[74] ibid art 7(3).
[75] European Interactive Digital Advertising Alliance, 'YourOnlineChoices.eu - About' <http://www.youronlinechoices.com/uk/about-behavioural-advertising> accessed 7 March 2017.

Google has expanded its empire, one such way is to merge various services into one single system. In 2012, Google announced the update of its uniform privacy policy that would cover all its major services, which allowed cross-service sharing of user profile data.[76] This aroused concerns over the compatibility with the principle of purpose limitation,[77] particularly due to the fact that users were not given the option to keep apart their profiles across those services. Data aggregation can be very powerful, especially when the datasets come from different contexts, and yet, there is no explicit regulation over such a practice in the GDPR. One point that might be relevant is hidden in Recital 43, which states that '[c]onsent is presumed not to be freely given if it does not allow separate consent to be given to different personal data processing operations'.[78] Strangely enough, however, this requirement has no reflection in the corresponding main-text Article. To minimise the threat of the network effect of data use to consumer choice, it could be a solution to allow data subjects the right — or even by default to require data controllers — to 'unbundle' the processing of data for multiple services, as well as their consent.

5.3.3 Necessity — The multiple dimensions of the necessity test

The last section on consent has touched upon an interesting point concerning the necessity test, which is worth a few more paragraphs here. As noted above, in the case of OBA, it is subject to constant debate whether certain practices of data processing are 'necessary' at all. The meaning or criteria of necessity are not given by the GDPR itself, and might be subject to change from context to

[76] , 'Google User Data to Be Merged Across All Sites Under Contentious Plan' *The Guardian* (25 January 2012) <https://www.theguardian.com/technology/2012/jan/25/google-merge-user-data-privacy> accessed 25 July 2017.
[77] Judith Rauhofer, 'Of Men and Mice: Should the EU Data Protection Authorities' Reaction to Google's New Privacy Policy Raise Concern for the Future of the Purpose Limitation Principle?' (2015) 1(1) European Data Protection Law Review 5.
[78] GDPR, art 43.

context. This also echoes the discussion in Chapter 4 where different senses of the necessity question have been identified: Is advertising necessary? If so, is OBA necessary? If so, is real-time-bidding-based OBA necessary? Of course, from a purely doctrinal point of view, consent given to advertising-related activities can be separate from that for the main service, and depending on the wording of the stated purpose ('optimised direct marketing', for example), the answers might be 'yes' to all three questions. Still, from a consumer protection point of view, just because something is necessary for a particular purpose does not mean that it is necessarily in line with consumer welfare.

That is not to say that 'direct marketing' is not a legitimate purpose itself. On the contrary, direct marketing is explicitly recognised as a legitimate purpose or interest by the GDPR.[79] However, while direct marketing may constitute a perfectly legitimate purpose, it is not necessarily as acceptable as other purposes. There are clearly some purposes that the GDPR considers more tolerable. Under Articles 5(1)(b) and 89, for example, processing for archiving, research or statistical purposes is generally not considered incompatible purposes, provided that appropriate safeguards are in place.[80] As for direct marketing, there are also provisions specifically dealing with data processing for this purpose. Article 21 provides that data subjects have the right to object to processing for direct marketing at any time, and the data controller may no longer process the data if such an objection is raised.[81] In the context of OBA, however, the provisions specially designed for direct marketing would have very little significance. As illustrated in the preceding chapter, since consent is the only legal basis for the use of cookies or similar techniques, which are critical for the operation of an OBA system, data

[79] ibid Recital 47.
[80] ibid arts 5(1)(b), 89.
[81] ibid arts 21(2), (3).

controllers would certainly need to obtain consent from data subjects. In that case, data subjects may always withdraw their consent at any time to achieve the same effect as exercising the right to object. Either way, the data subject may stop the processing of data but at the cost of losing the service.

Therefore, for Internet users, although direct marketing is treated as a special purpose of data processing, it does not make a difference in terms of the level of protection. Yet, that purpose is special for a reason. Direct marketing is among those purposes that benefit mostly — if not entirely — the data controller. As such, in cases where the data subject gives consent to processing for a purpose in the interest of the data controller, a higher level of protection — for example, by applying a stricter test of necessity — should be afforded. Furthermore, even within the spectrum of direct marketing, the degree of power asymmetries and level of intrusion may vary, depending on the technical scheme and business model. In this regard, the criterion of 'necessary' in the GDPR — whether in the form of the data minimisation principle, the other legal bases other than consent, or the supportive complement to existing safeguards — fails to mirror such diversity. It neither distinguishes different purposes nor gives adequate consideration to other factors that might counterbalance the necessity element. What is imposed is a one-size-fits-all standard. The bottom line is that, a range of different combinations of safeguards, as will be explained in the next chapter, should be available in proportion to the sensitivity of the purpose.

5.4 Guaranteeing meaningful participation and informed decision-making in political contexts

5.4.1 The blurred line between private and public use of personal data

The last decade has witnessed some dramatic changes in the field of use of personal data, and the line between the private and public sectors in terms of

data use has become much less distinctive. The vanishing boundaries between 'private' and 'public' are evident in two slightly different senses: First, personal data now moves freely through a 'revolving door' between public and private bodies.[82] Second, even with personal data utilised completely by private entities, there could be profound implications for a society's public life. The political implications of the ubiquitous presence of OBA might represent a case in point. As summarised in Section 3.5, the potential adverse effects of OBA on democracy are three-fold: It might lead to the chilling effect; it might cause filter bubbles; and it might give political campaigns unfair advantage.

The public implications of private use of OBA data are now even more evident, for example with the Cambridge Analytica scandal, where a privately-owned, non-political-party entity may exert political power. It is clear that personal data collected by players in the private sector can be easily used for political micro-targeting. Irrespective of the actual powerfulness of such practices in manipulating public opinions, the employment of such techniques itself is likely to be incompatible with the idea of democracy.[83] In such a context, what is at stake is a special category of interpersonal interest: political interest. This interest represents a value that is unique in three dimensions. First, unlike the commercial interests discussed in the previous section, political interest concerns public affairs and public life. Second, and consequently, the interest may be claimed vis-à-vis not just a certain group of actors — whether private (businesses) or public (government) in natural —

[82] Judith Rauhofer, 'Round and Round the Garden? Big Data, Small Government and the Balance of Power in the Information Age' (2014) 2014/06 University of Edinburgh School of Law Research Paper Series <http://ssrn.com/abstract=2389981> accessed 22 October 2014.
[83] For different opinions on this matter, see Balázs Bodó, Natali Helberger and Claes H. De Vreese, 'Political Micro-targeting: A Manchurian Candidate or Just a Dark Horse?' (2017) 6(4) Internet Policy Review 3, 4-5.

but indeed the entire society. Third, this interest implies not only rights but also duties assigned to members of the community.

In a progressive democracy, data subjects are regarded as important participants in the course of the governance of public affairs, or citizens. Since citizen is an important aspect of a data subject, an effective data protection legal framework cannot be considered sufficient without factoring in the political effects of use of personal data. Many specific points discussed in previous sections are applicable to this context. Individual autonomy, for example, certainly involves participation in political life, as '[d]enying one the ability to engage in [it] curtails to a significant degree one's ability […] to feel a full member of a political community.'[84] Also, the imbalance of power in an economic setting is largely applicable to the political one. However, this section will mainly focus on those issues that are specifically relevant to a political context.

5.4.2 Consent — Individual decision vs common future

The sensitivity of political use of personal data is not completely ignored by the data protection legal framework. Both the DPD and the GDPR categorise personal data revealing 'political opinions' as sensitive data and impose additional restrictions on the processing of such data.[85] For instance, processing of sensitive data is prohibited as a matter of principle, although a number of exceptions, including explicit consent, are also provided by both instruments.[86] Member States may also overrule the consent exception, deciding that the prohibition cannot be waived by means of consent.[87] Such a

[84] Raz, *The Morality of Freedom* (n 4) 410.
[85] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281/31 ('DPD'), 8(1); GDPR, 9(1).
[86] DPD, art 8(2)(a); GDPR, art 9(2)(a).
[87] DPD, art 8(2)(a); GDPR, art 9(2)(a).

restriction on processing personal data with political implications is necessary but far from adequate.

The problem is that while some activities of political targeting are based on data about personal political views,[88] not all such targeting services require personal data revealing 'political opinions'. Deutsche Post, for example, was revealed to have sold household data to two German political parties for targeting purpose.[89] Although the company argued that the dataset did not contain sensitive data — and implied that it was integrated, not personal data at all — concerns over the ethics of such practices remained.[90] The nature of online behavioural data, especially social media data, can be even more complicated in that, while not all data necessarily indicates the data subject's political view, certain inclinations might well be inferred from such data. For example, in the case of Cambridge Analytica, while the demographic data, behavioural data and user-generated content is not necessarily associated with certain political opinions directly, a psychometric profile inferred from such data can be good indicators of certain categories of public attitudes.[91] Arguably, these profiles constitute sensitive data that reveals data subjects' political opinions.

---

[88] This is particularly the case when personal data is provided by political parties themselves. It is reported that, for example, in the US presidential elections, both political parties operate massive voter databases to build revealing voter profiles. For the details, see Daniel Kreiss, *Prototype Politics: Technology-intensive Campaigning and the Data of Democracy* (Oxford University Press 2016) 204-220.

[89] Deutsche Welle, 'Deutsche Post Defends Voter-microtargeting Data Practice' *Deutsche Welle* (2 April 2018) <http://www.dw.com/en/deutsche-post-defends-voter-microtargeting-data-practice/a-43223747> accessed 8 April 2018.

[90] ibid.

[91] Frederike Kaltheuner, 'Cambridge Analytica Explained: Data and Elections' *Medium* (13 April 2017) <https://medium.com/privacy-international/cambridge-analytica-explained-data-and-elections-6d4e06549491> accessed 1 March 2018; Jonathan Albright, 'Cambridge Analytica: the Geotargeting and Emotional Data Mining Scripts' (2017) <https://medium.com/tow-center/cambridge-analytica-the-geotargeting-and-emotional-data-mining-scripts-bcc3c428d77f> accessed 20 November 2017.

More importantly, in theory, effective political targeting sometimes does not have to involve political data. As mentioned above, Facebook itself runs a service facilitating political campaigns,[92] and this service, according to a study, is based on 'location, age, education, gender, and other demographic feature' only.[93] However, just because political data is not involved does not mean that such activities have no political implications. The use of non-political data can nevertheless result in filter bubble effect and unjust competitive edge. In terms of the filter bubble effect, personalisation of news presenting does not have to be able to ascertain or predict the data subject's political preference. Instead, this can be done solely based on demographic or online behavioural data. As for the unjust competitive edge that some political campaigners might gain, they do not have to know about people's political attitudes. All they need to know is probably the voter's neighbourhood, consumption pattern or lifestyle. This would nevertheless allow them to package their political message in a fashion appealing to individual voters. The problem here lies in that it is the political nature of the purpose, not the political nature of the data itself that should be addressed. In other words, political marketing could function perfectly well without political personal data at all, unless we redefine what 'political personal data' is. Redefining political data, however, may well end up turning all data into political data as long as such data can be used for political targeting. It seems that it is not the nature of the data that matters as much as the nature of the purpose.[94]

---

[92] Facebook, 'Facebook Elections' <https://www.facebook.com/business/a/politics-industry> accessed 27 June 2016.

[93] Jessica Baldwin-Philippi, 'The Myths of Data-Driven Campaigning' (2017) 34 Political Communication 627.

[94] Yoan Hermstrüwer, 'Contracting Around Privacy: The (Behavioral) Law and Economics of Consent and Big Data' (2017) 7(3) JIPITEC 9, 15.

A remark similar to the conclusion of the last section can be made here: Not all purposes are created equal. For those that would cause negative externalities [95] or pose threat to the public good of a society, [96] further restrictions should be imposed. Unfortunately, data processing for political purposes has no special status in the GDPR. It should be subject to a stricter approach to consent, or closer oversight by independent authorities. One might even take a further step to argue that the use of personal data for political campaigns cannot be based on consent, since ethically speaking, individuals have no moral right to consent to activities that would erode the collective democracy of the society.[97] That is exactly why a large part of public discussion following the Cambridge Analytica revelations might have been misplaced. The focus is largely on the fact that Cambridge Analytica harvested massive amounts of data from a third-party app without user consent.[98] While this is indeed an important legal point, what has been missed here is the following question: If valid consent had been secured, does that mean they should be allowed to do so? In any event, data protection law should lay down

[95] Mark MacCarthy, 'New Directions in Privacy: Disclosure, Unfairness and Externalities' (2011) 6(3) I/S: A Journal of Law and Policy 425; Hermstrüwer (n 94).
[96] Priscilla M. Regan, *Legislating Privacy: Technology, Social Values, and Public Policy* (University of North Carolina Press 1995) 225-227; Joshua A.T. Fairfield and Christoph Engel, 'Privacy as a Public Good' (2015) 65 Duke Law Journal; Mireille Hildebrandt, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology* (Edward Elgar 2015).
[97] See Robin West, 'Authority, Autonomy, and Choice: The Role of Consent in the Moral and Political Visions of Franz Kafka and Richard Posner' (1985) 99(2) Harvard Law Review 384, 424.
[98] See The Guardian, 'The Guardian View on Data Protection: Informed Consent Needed' *The Guardian* (19 March 2018) <https://www.theguardian.com/commentisfree/2018/mar/19/the-guardian-view-on-data-protection-informed-consent-needed> accessed 8 April 2018; Hilary Osborne and Jessica Elgot, 'MPs Summon Mark Zuckerberg, Saying Facebook Misled Them' *The Guardian* (21 March 2018) <https://www.theguardian.com/uk-news/2018/mar/20/officials-seek-warrant-to-enter-cambridge-analytica-hq> accessed 8 April 2018; Julia Carrie Wong, 'Facebook's Privacy Practices are Under Investigation, FTC Confirms' *The Guardian* (26 March 2018) <https://www.theguardian.com/technology/2018/mar/26/facebook-data-privacy-cambridge-analytica-investigation-ftc-latest> accessed 8 April 2018.

certain restrictive safeguards with regard to consent in the context of political targeting, which, unfortunately, is nowhere to be found in the GDPR.

5.4.3 Necessity — Or purpose desirability?

This leads to a relevant question concerning the ethics of political targeting through OBA. To be precise, the question concerns whether political targeting should be allowed as a legitimate purpose itself.[99] Again, the GDPR has not shed any light to this matter and in practice, OBA is actually being used by many political campaigns around the world.[100] Legally speaking, if such processing is held illegitimate as a purpose for data processing, then such practices will be entirely disallowed under the GDPR as it would be an infringement on the purpose limitation principle.

Before the arrival of the big data trend, political targeting was possible but done in a much less effective, sophisticated manner. It is one thing to put flyers into the mailboxes in particular neighbourhoods, and quite another to use algorithms to personalise a campaigning video based on user profiles on social media and see how it is received across constituencies. Again, the extent to which online political targeting can be done is a matter for public debate, although it is unlikely that it would be considered entirely illegitimate. In any case, it should be acknowledged that political marketing is much more sensitive than a lot of other purposes. The GDPR, as it currently stands, offers

---

[99] For the latest development of the public discussion about the ethics of political targeting based on personal data, in particular in the UK, see Carole Cadwalladr, 'The Great British Brexit Robbery: How Our Democracy Was Hijacked ' *The Guardian* (7 May 2017) <https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy> accessed 25 July 2017; Robert Booth, 'Inquiry Launched into Targeting of UK Voters Through Social Media ' *The Guardian* (17 May 2017) <https://www.theguardian.com/technology/2017/may/17/inquiry-launched-into-how-uk-parties-target-voters-through-social-media> accessed 25 July 2017.

[100] Google and Facebook have marketing services that are specially designed for political campaigns. See Facebook (n 92); Google, 'Win the Moments That Win Elections' <https://www.google.com/ads/elections/> accessed 7 January 2017.

only two options for the fate of a political campaign: fully legitimate or fully illegitimate. Once again, data controllers and data subjects are left with only two options: either an overall prohibition of such processing, or permission to such data processing just like other purposes. In reality, the case might be that the purpose of political targeting is not illegitimate itself, but it is so risky that further safeguards are needed. The lack of something like a restricted zone makes it hard to allow at least some forms of political targeting where the risks are relatively low but additional safeguards are needed.

The underlying problem here is the conceptual structure of the necessity test. The requirement of necessity has pervaded the entire GDPR, not just as a principle itself (data minimisation), but also underpinning all of the legal grounds for lawful data processing (including even consent). However, the necessity test cares only about the connection between the processing in question and its purpose. It does not take one step further to question the 'necessity' of the purpose. The word 'necessity' is put in quotation marks here because maybe necessity is not the right question to ask. Of course, some purposes have their second-order purposes and would therefore be worth questioning whether such purposes are necessary for their second-order purposes. Yet, in many cases, the purpose concerned does not have an additional, clearly-defined purpose behind it. In that case, what needs to be further investigated would be the desirability or acceptability of these purposes. In other words, even if certain forms of data processing prove to be necessary for political targeting, what follows should be more questions concerning the necessity and desirability of political targeting itself. As underlined in Section 5.3 in the context of consumer protection, the GDPR has clearly distinguished certain types of purpose and has adopted special rules for them, with direct marketing being one of them. Even leaving aside the

practical utility of those special rules, the concept of direct marketing is such a broad one itself that it encompasses both benign and malign instances. If political direct marketing belongs to the latter category, data protection law should take additional care of it.

5.5 Summary: Moving forward but lagging behind

5.5.1 Policy choice and autonomous, interpersonal and political interests

In this chapter, the effectiveness of the GDPR in protecting data subjects has been examined through three lenses: autonomous, interpersonal and political interests. These three categories of interest are not chosen arbitrarily; instead, they signify the subtle differences in a series of variables in several continuums. The first dimension to see the differences is the nature of such interests. Obviously, the discussion surrounding autonomy has more to do with individualist values, namely informational self-determination, liberty and equality. The commercial, interpersonal interest mirrors the economic welfare of a collective of individuals. The political interest signifies the democratic value, which is the constitutional legitimacy of the society.

The distinction between these interests leads to a second dimension of observations: the role of social preferences in the course of identifying the appropriate approach to realising such interests — which entails further explanation here. There are many things in life to which the preference of the society does not matter at all. For instance, in a secular nation, just because the majority of the population shares strong piety to a particular religious faith would not justify the state's attempt to force the non-believers to follow that religion. However, on other matters, it is reasonable to impose social consensus on individuals, irrespective of their own individual preferences. In many countries, for example, pension schemes are mandatory and the age at

which one may begin to receive pensions is provided by law. This is not the place to explain the different rationales behind such decisions, but the tension between personal and social choice is evident in almost all aspects of life. Such a tension exists between how different categories of interests are protected. For fundamental rights, in particular the individualistic ones, the exercise of them is usually considered a matter of personal choice, and thus personal preference often takes precedence. In the case of commercial interests, social preference might account for a larger part in decision-making, partly because economic efficiency often requires collective coordination, and also partly because individuals (like consumers) are often in a relatively weaker position if they do not take collective initiatives. As regards constitutional legitimacy, social preference would play a dominant role as individuals' rights (and even obligations) to participate in public governance would affect all members of the society.

The degree of the impact that a social preference has would largely determine the third dimension: policy choice. Where personal preference has the priority, individual choice should be given utmost respect. Accordingly, regulatory intervention should generally be minimised, and mutual agreement should be respected as far as possible. Where there is a compelling social preference, it would be justifiable to impose certain regulatory constraints, and even overriding any personal choice in the form of mutual agreement.[101] If there is a dominant social preference that would yield desirable outcomes, but without a compelling justification, the policy choice may be a hybrid of both individual choice and regulatory constraints. This can be achieved, for example, by re-arranging legal, social, technological or

[101] Ian Brown and Christopher T. Marsden, *Regulating Code : Good Governance and Better Regulation in the Information Age* (The MIT Press 2013) 63-68.

economic configurations in which personal choice can be shaped, but not determined, towards a given direction. The case of organ donation serves as an interesting illustration of these possibilities. It is not surprising that organ donation is universally implemented on a voluntary basis, but the way potential donors express their will varies greatly from one culture to another. In some European countries, cadaveric organ procurement is carried out based on 'presumed consent', which means a deceased person is considered to have consented the donation, unless they have explicitly opposed before death.[102] Unsurprisingly, this has a positive effect on organ donation rates.[103] In the context of data protection, the implication is that, depending on the interest that is meant to be protected, there are more policy options than what has been employed in the GDPR.

Interest	*Personal/social preference*		*Policy choice*
Personal autonomy	Compelling personal preference		Individual choice
Commercial interests	Dominant and desirable social preference		Hybrid
Collective values	Compelling social preference		Regulatory constraints

Table 5.1: Various dimensions concerning the protection of data subjects

Two points need to be stressed here: The relationship between the dominance of social preference and the appropriate policy choice is not absolute. In other words, the policy options should be kept open to various configurations of personal/social preference. Additional considerations, as will be further discussed in next chapter, could have impact on the final choice between policy options. Second, social preference here is an institutional, theoretical

[102] Eric J. Johnson and Daniel Goldstein, 'Do Defaults Save Lives?' (2003) 302 Science 1338; Alberto Abadie and Sebastien Gay, 'The Impact of Presumed Consent Legislation on Cadaveric Organ Donation: A Cross-country Study' (2006) 25(4) Journal of Health Economics.

[103] Abadie and Gay (n 102).

concept, and is not necessarily identical to what the majority members of a community actually prefer. For instance, if the legislature of a jurisdiction conscientiously decides that a regulatory restriction is necessary for the fundamental interests of the people, their decision could be justifiable despite the fact that, say, opinion polls suggest otherwise.

5.5.2 How the GDPR falls short of its objective

By analysing the multi-dimensional characteristics of various types of interests, the shortcoming of the regulatory model of the GDPR can be better seen now: The protean roles that data subjects play today call for a more versatile, responsive legal framework. The problem of the 'consent + necessity 2.0' model is that both elements suffer a high degree of inflexibility in themselves, leaving no space for a potential third approach to develop. Without such a transitional territory between those two options, the data protection framework would neither satisfy the needs of data subjects nor those of data controllers — some would argue the minimum requirements are not broad enough, others would argue individual choice should be applied to more areas. The 'consent + necessity 2.0' model shows a perpetuated 'libertarianism-paternalism' conceptual binary, and fails to realise the complex interests that data subjects enjoy may be achieved through a spectrum of regulatory measures.

Apart from this general flaw, specific shortcomings regarding consent and necessity have also been identified. As alert readers might have noticed, what these distinctive roles demand is not always aligned, and sometimes even conflicting with one another. For instance, when data subjects are put in the light of individual autonomy, much more emphasis has been placed on the effective exertion of consent. However, when data subjects are considered citizens in a democratic society whose informed participation in public affairs

deserves protection, the significance of state interference has been stressed. These distinctive features, as pointed out above, stem from the different nature of the interests these capacities represent, and point to slightly different problems that the GDPR suffers.

Interest	*Consent*	*Necessity*
Autonomous	Rigidity and flatness	Inability to signify priority and seriousness
Interpersonal (Commercial)	Failure to diversify choice	Inability to capture the uniqueness of the 'direct marketing' purpose
Political	Misplacement of focus on the nature of data	Insufficient categorisation of sensitivity of purpose

Table 5.2: The shortcomings of the 'consent + necessity 2.0' model

In terms of consent, the GDPR has largely retained the conceptual model developed in the age of the DPD. It has not mirrored the demands and the possibilities of a revolutionised concept of consent enabled by big data. A more delicate, customisable approach to consent may be what we need in the age of big data. As regards necessity, it is also crucial to bring in various levels of necessity as well as various levels of desirability. To tackle the 'take-it-or-leave-it' challenge, a data protection regime could have done more to bridge the gap between the 'take it' and 'leave it' options. The point of such a bridge is to create incentives for those who give consent to everything to start realising that they have more alternative choices other than cutting off entirely from the Internet. A valid criticism, however, can be that it might also 'lure' those tough-minded on privacy into something much softer. It remains to be seen whether the majority of the society would actually move up or down the data protection ladder.

Of course, there are a lot of other uncertainties surrounding these suggestions. Each of them marks most likely a beginning of a new strand of debate that is far from a conclusive answer. Still, it is important that we are able to spot both the strengths and weaknesses of the current data protection

framework. Even more importantly, we need to realise that, while big data has posed and will probably continue to pose more challenges to the values that we as a society cherish, it also offers new possibilities to strengthen existing safeguards and to develop new ones. All new proposals — technical, economic or legal — will undoubtedly have to be examined in greater depth both theoretically and practically. There is no guarantee that all new ideas will develop into concrete plans that are feasible and effective, but they should be included as part of the ongoing discussion of how to improve our data protection law. Some of such new concepts and possibilities have been brought up lightly throughout this chapter. In the light of the current legal framework's weaknesses as exposed above, they will be further explored and elaborated in the next chapter.

Chapter 6 Constructing a Big Regulatory Toolbox for Big Data: Exploring Alternative Approaches of Data Protection Law in the Context of Online Behavioural Advertising

A key message conveyed in the last chapter is that, in a highly complex digital world where a diversity of technological and social conditions are reshaping the dynamics of interests and values, the 'consent + necessity 2.0' model featured in the current data protection paradigm is struggling to adapt to the techno-social realities. The GDPR represents a laudable milestone in that it has strengthened many dimensions of the DPD to make it harder to exploit the loopholes. However, as pointed out above, the binary model has neither fully captured the heterogeneous nature of the spectrum of issues at hand, nor has it fully realised the potential solutions enabled by a hybrid of legal, technological and economic mechanisms.

There is therefore a pressing need to review the general regulatory approach to big data, which requires both a re-discovery of the regulatory measures already in place in the existing legal framework, and also a fresh perception of how new instruments may be introduced so as to equip regulators with a better toolbox. It might be overly ambitious, and premature, for this study to come to conclude whether and how specific tools should be utilised in a given context, but at least, it would be plausible to discover a range of such tools, and to locate where each of them fits into the broader picture of a holistic strategy to address the big data issues.

As in previous chapters, the enquiry of this chapter will again be informed by the case study of OBA, based on the legal framework of the GDPR as

extracted in Chapter 4. Given the breadth of matters covered by the GDPR, it makes sense to keep the discussion on a relatively narrow perspective. In the course of revealing the possible instruments in the regulatory toolbox, the discussion in this chapter will only focus on the authorisation scheme for processing of personal data, namely the synthesised mechanisms whereby certain categories of data use are allowed or prohibited. This means that the more procedural or supportive roles of data protection law, such as enforcement, education and remedies, will be largely set aside. Such an approach resonates with the one taken in previous chapters as the discussions on the trade-offs made in the context of different data uses (Chapters 2-3) and the legal regime characterised by the 'consent + necessity 2.0' model (Chapters 4-5) have been framed mostly with regard to the conditions under which data processing may be permitted. To such an extent, this chapter will also serve to recollect the findings offered in previous chapters, and connect all the dots throughout the journey of investigation.

6.1 The black and white of data protection

6.1.1 The black- and whitelist native to data protection law

(a) Deconstructing the authorisation scheme as lists

One of the core functions of data protection law is to lay down clear rules whereby it be can decided whether a given operation of data processing may take place or not. The entirety of such rules under a particular data protection regime operates as an authorisation scheme. Such a scheme provides a set of rules whereby one may determine, based on the circumstances surrounding the processing of personal data in question, whether the processing can be allowed or not. To the extent that it is possible to compile all the factors to be considered under a data protection regime into a permissive list and a

prohibitive list, an authorisation scheme can be viewed as made up of a whitelist and a blacklist.

Deconstructing the authorisation scheme as lists is not just about the choice of a metaphorical presentation, but is also of epistemic importance as it provides a helpful angle to examine a data protection regime, both doctrinally and critically. In the rest of this chapter, it will be demonstrated that, first, data protection law *can be* conceptually viewed as functioning by specifying the realms and mechanisms of those permissive and prohibitive lists; second, data protection law *should be* viewed this way as it offers data protection policymakers and regulators a better handle on the possible regulatory initiatives; third, data protection policymakers and regulators *may* (and *should*) in practice consider codifying certain data processing practices into such lists, based on well-grounded justifications and subject to constitutional constraints. These ideas are hereinafter referred to as a 'list-based' approach to data protection law, which will be explained in greater detail along with the analysis that will extract the white- and blacklist from the GDPR.

(b) The built-in whitelist in the GDPR

The starting point of understanding EU data protection law, which may fundamentally differ from a regime in another jurisdiction, is that processing of personal data is generally forbidden unless specifically permitted by law. Having said that, the current legal framework has provided a rather broad scope of rules under which data processing is permitted. These rules, as will be shown below, are scattered over various parts of the GDPR and operate within a rather complex hierarchy. By mapping out all these permissive rules that have been readily built into the architecture of the GDPR, it is possible to draw up a data processing whitelist.

In the first place, there are certain exemptions to the material scope of the GDPR under which data processing of such kinds would not be subject to EU data protection law.[1] These exemptions include, *inter alia*, purely personal or household activities[2] and matters in criminal procedures.[3] To the extent that those activities would not be restricted in the absence of further national regulation, they can be considered part of the whitelist. Apart from these generic exemptions, the most common whitelist scenarios take the form of legal grounds provided by Article 6 (for general personal data) and Article 9 (for sensitive personal data). Points (a) to (f) of Article 6(1) — regarding consent, contract, legal obligation, vital interests of natural persons, public interests and legitimate interests — each essentially provide a whitelist measure. The same goes for Article 9(2), where the legal grounds for processing of sensitive data are provided for lawful processing of sensitive data. Another whitelist exemption is hidden in the principle of purpose limitation. As explained in Chapter 4, with the interpretation of Recital 50, purpose limitation in effect allows further processing of personal data for certain purposes that are deemed to be compatible with the original ones, with no need for a separate legal basis.[4] That means, such compatible further processing is allowed even in the absence of a legal basis (and thus 'whitelisted'), as long as the initial collection of data is justified on one of those bases.

Despite all these whitelist exemptions by which data processing is allowed, most of them are hardly applicable to the case of OBA, or generally speaking,

---

[1] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1 ('GDPR'), Recital 112.
[2] ibid art 2(2)(c).
[3] ibid art 2(2)(d).
[4] ibid art 5(1)(b), Recital 50.

commercial use of big data. Apart from individual consent, these whitelist options mainly concern either highly public objectives, such as public interest[5] and public tasks[6], or highly personal matters, such as purely household activities[7] or vital individual interest[8]. It is almost impossible to reasonably invoke most of these legal bases to justify commercial uses of personal data. This is consistent with the observation in Chapter 4 that the only viable legal grounds for marketing purposes would be consent, contract and legitimate interest. It is however also concluded later in the same chapter that contract and legitimate interest would not form a practical basis for the typical operation of OBA systems as it is highly difficult to prove the activities involved are indeed necessary for the performance of the main contract or the legitimate interest of the marketers.

(c) The built-in blacklist of the GDPR

Even where allowed by a whitelist provision, it is possible that the authorisation of a data processing operation may be overridden by a prohibitive rule. Throughout the GDPR, an array of restrictions and exceptions have been set out to de-legitimise certain data processing activities under specific circumstances. These further restrictions, including those generally applicable to all legal bases and those specifically applicable to a particular legal basis, essentially function as a blacklist that prohibits certain types of data processing even where a legal ground exists.

For example, the data protection principles can be largely seen as a general set of criteria to preclude data processing contrary to the fundamental consensus on how personal data should be treated, no matter whether and on

[5] ibid art 6(1)(e).
[6] ibid Recital 112.
[7] ibid art 2(2)(c).
[8] ibid art 6(1)(d).

which basis such processing is specifically justified. It bears repeating that the violation of any principle provided by Article 5, unless within the limited scope of exemptions,[9] would result in the processing in question being unlawful, even if a legitimate basis has been established. For example, if the data processing in question is found unfair under the principle of fairness,[10] it would be prohibited regardless of the existence of a legal basis. Likewise, any processing that is not necessary for the purpose (thus against the data minimisation principle)[11] or serves an illegitimate purpose (against the purpose limitation principle)[12] would not be permitted. As such, these principles constitute the most abstract layer of the data protection blacklist.

Additionally, some categories of data processing operations are subject to further restrictions, again, irrespective of the legal basis chosen. For instance, as analysed in Chapter 4, automated individual decision-making, including profiling, is generally prohibited under Article 22 unless one of the three exceptions (contract, explicit consent and authorisation by law) applies.[13] Another example is the prohibition imposed by Article 44 on data transfers to territories or international organisations outside the EU/EEA ('third countries').[14] Such transfers may take place only if the third country is recognised by an adequacy decision,[15] or appropriate safeguards are in place,[16] or one of the derogations applies.[17] Those derogations include, *inter alia*, explicit consent, contract, public interest, etc.[18] To the extent that data

[9] ibid art 23.
[10] ibid art 5(1)(a).
[11] ibid art 5(1)(c).
[12] ibid art 5(1)(b).
[13] ibid art 22.
[14] ibid art 44.
[15] ibid art 45.
[16] ibid art 46.
[17] ibid art 49.
[18] ibid.

processing covered by either Article 22 or 44 is subject to a prohibition in principle, these two Articles have created an extra layer of blacklist for processing activities that would be otherwise allowed based on one of the grounds provided by Article 6. It should however be noted that the derogations included in Articles 22 and 44 mean that a further whitelist is also operational against this blacklist, thereby establishing a rather complex 'rule—exemption—exemption from exemption' logical structure. Having said that, this does not change the general nature of these Articles being blacklist measures in that additional safeguards are required in either case,[19] where a failure to comply with these safeguards would render the processing unlawful even if explicit consent has been obtained.

There is a third layer of blacklist measures applicable to specific grounds covered by Article 6 (personal data of general nature) and Article 9 (sensitive data). These measures usually take the form of additional exceptions or conditions. The 'legitimate interest' ground provided by Article 6(1)(f), for example, applies 'except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject'.[20] In other words, even where the processing concerned is proved necessary for a legitimate interest, it would nevertheless be prohibited if it fails the balancing test against the interests of the data subject. Similar thresholds can be found in, among others, Article 6(3) (the proportionality requirement applying to the grounds of legal obligation and public interest), Article 9(2)(d) (the conditions for processing of

[19] In the case of automated individual decision-making, such safeguards include 'at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision' (Article 22(3)), and in the case of data transfers to third countries, 'activities carried out by public authorities in the exercise of their public powers' (Article 49(4)) or 'limits to the transfer of specific categories of personal data to a third country or an international organisation' imposed by law (Article 49(5)) cannot be exempt by means of explicit consent.

[20] GDPR, art 6(1)(f).

sensitive data by political, philosophical, religious organisations or trade unions), Article 9(2)(g) (a proportionality requirement upon processing of sensitive data for the purpose of public interest), Article 9(2)(h) (professional secrecy requirement for processing of sensitive data for the purposes of preventive or occupational medicine, read together with Article 9(3)), Article 9(2)(i) (suitable and specific measures for processing of sensitive data for reasons of public health), Article 9(2)(j) (the proportionality test for processing of sensitive data for archiving, research or statistical purposes), Article 9(4) (national limitations on processing of health-related data) and Article 21 (cessation of processing on grounds of public interest and legitimate interest in the case of objection).

A blacklist measure can also preclude the effect of a mutual agreement between the data subject and data controller. A typical preclusion of this kind can be found in Article 7(4): As discussed in detail in Chapter 4, consent bundled with the provision of service is presumed to be not freely given and hence invalid. Article 8 also imposes an age limitation on children's consent in relation to information society services. As for sensitive data, Article 9(2)(a) contains a proviso whereby EU or national law may uphold the prohibition with regard to sensitive data regardless of explicit consent given by the data subject.[21] This provision is a continuation of Article 8(2)(a) of the DPD but rarely have Member States invoked this article to create a national blacklist.[22] A similar specification is included in Article 49, which precludes consensual agreement as a legal basis for transfers of personal data to third countries where the processing is carried out by public authorities.[23] By explicitly

[21] ibid art 9(2)(a).
[22] Douwe Korff, *Data Protection Law in Practice in the European Union* (Federation of European Direct and Interactive Marketing and The Direct Marketing Association 2005) 46.
[23] GDPR, art 49(3).

overriding individual decisions to authorise certain categories of data processing, these measures in effect already function as a blacklist that counteracts the individual decisions to authorise data uses by means of, say, their consent. The implications of this are twofold: On the one hand, consent is not a panacea for any form of data use; rather, certain restrictions have been imposed on the use of consent as a legal basis. On the other hand, such restrictions are limited to a very narrow scope. As will be shown below, the list of restrictions on individual consent may be, and sometimes even should be, expanded where certain individualistic, collective and public interests are gravely endangered.

6.1.2 The expandability of the white- and blacklist

(a) Expanding the whitelist

The brief survey above of the white- and blacklist measures readily in place in the GDPR reveals that the authorisation scheme under the current data protection law largely operates between the whitelist and blacklist measures. The decision to include a particular category of data processing into a white- or blacklist reflects the policymakers' value judgment on the nature of the kind of data processing in question. The problem is, on what basis can policymakers defend such decisions? What can they do if they have identified further types of data uses that they consider appropriate on the white- or blacklist? Are these further measures subject to any limitation or scrutiny? This section is intended to address these questions by exploring the potential mechanisms by which the white- and blacklist can be expanded, either by the policymakers or the regulators. Again, the enquiry will begin with the whitelist measures.

As pointed out in Chapter 2, while most of the arguments advanced by the marketing sector in favour of uses of personal data for OBA purposes do not hold water, there remain certain valid points. The plausible arguments include

part of the 'free content and services' and 'boosting the digital market' arguments, although both are often exaggerated. The next question then is whether and how the legitimate parts of these potential benefits can be mirrored in certain types of data uses that can arguably be whitelisted. Perhaps drawing up such a whitelist would prove immensely challenging in the context of online marketing, because there do not seem to be any interests in this sector that are compelling enough to go beyond consent as it currently stands in the GDPR. If it is true that 'free content and service' is all about choice[24] and 'boosting the digital market' is all about trust,[25] it would be extremely unreasonable to force such policy goals on individuals by making some categories of data use lawful independent of individual choice — it simply runs counter to the nature of either goal.

That said, the possibility to include a limited number of OBA activities into the whitelist cannot be fully ruled out at this point. Also, in the wider context of big data there may well be a multitude of reasons why certain data processing activities should be whitelisted. One potential claim supporting such an inclusion may, for instance, rest in the free-rider problem. It can be argued that, for example, Internet users who have chosen not to share any personal data to the service provider are actually taking a free-ride in particular when the service is fully financed by advertising.[26] For service providers who can hardly change to a paid-for or freemium model, one may

[24] Commission, 'A Digital Single Market Strategy for Europe' (2015) Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions COM(2015) 192 final, 3.
[25] ibid 9.
[26] See Mike Hammock and Paul H. Rubin, 'Applications Want to Be Free: Privacy Against Information' (2011) 17-18 <http://ssrn.com/abstract=1781906> accessed 3 May 2018; John Naughton, 'The Rise of Ad-blocking Could Herald the End of the Free Internet' *The Guardian* (27 September 2015) <https://www.theguardian.com/commentisfree/2015/sep/27/ad-blocking-herald-end-of-free-internet-ios9-apple> accessed 3 May 2018.

reasonably advocate keeping legitimate data uses on the whitelist. One of the ongoing discussions about the overhaul of the ePrivacy Directive concerns the introduction of 'web audience measuring' as a new exception to consent for use of cookies.[27] Such techniques are usually considered useful and low-risk as it does not target individuals.[28] Although the use of data for marketing purposes is significantly different from those merely for measuring purposes, the point here is that personal choice does not always defeat legitimate interest. The data protection whitelist is sometimes a useful regulatory tool in certain contexts where the use of data is beneficial to the majority of the society but vulnerable to individual decisions.

(b) Expanding the blacklist

The analysis above of the blacklist elements dispersed throughout the GDPR has shown that the prohibitive measures manifest themselves mainly on three threads in data protection law: (1) in the form of data protection principles; (2) in the form of restrictions on certain activities regardless of the legal basis; (3) in the form of additional conditions on specific legal grounds that would otherwise authorise the use of personal data. For data protection policymakers, such a perception leads to three major implications. First, one of — if not the most important of — the underlying logics of data protection law is to protect personal data by imposing prohibition on processing. Second, and as a result, a large part of the EU data protection regime is already substantially no different from drawing up a blacklist, or a number of blacklists, in that it

[27] European Parliament, 'Report on the proposal for a regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications)' (2017) A8-0324/2017, 58; Article 29 Data Protection Working Party, 'Opinion 04/2012 on Cookie Consent Exemption' (2012) 00879/12/EN WP 194, 11.

[28] European Parliament (n 27) 24.

outlaws certain types of data uses including those based on the data subject's consent, although formally those uses do not appear as a list. Third, and perhaps most importantly, given that the blacklist approach is already an option in the policymaker's regulatory toolbox, there is no doctrinal reason why its use cannot be expanded where this is reasonable to adapt to the changing social, economic, political or technological circumstances. In short, blacklisting certain types of data processing is nothing new in data protection law. Moreover, this approach is also common in other areas of law that aim at the protection of individuals as citizens or consumers, often in an even more 'list-looking' manner. In consumer protection law, for example, the Unfair Commercial Practices Directive includes an annex specifying the 'commercial practices which are in all circumstances considered unfair'.[29] A similar list of unfair contract terms is also annexed to the Unfair Contract Terms Directive,[30] under which the inclusion of those unfair terms would render a contract unenforceable.[31]

In data protection law, with the three layers of blacklist measures in mind, there are a number of legislative strategies available to policymakers if they wish to expand the data protection blacklist. The first option is, of course, putting in place new data protection principles or making existing principles more restrictive. However, this is probably the most challenging option given the overarching effects as well as the pragmatic need to keep such principles

[29] Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council ('Unfair Commercial Practices Directive') [2005] OJ L149/22 ('UCPD'), Annex I.
[30] Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts [1993] OJ L95/29 ('UCTD'), Annex I.
[31] GDPR, art 6(1).

stable. A second option would be to identify those areas that are particularly high-risk and then impose a blanket prohibition on related data processing activities. This is a more precise and practical approach, which would probably target the risky scenarios better, but overriding all legitimatising grounds may well turn out controversial for its potential spillover effects on legitimate uses of personal data. Stability would of course be another issue depending on how frequently the expanded blacklist is subject to change. Perhaps the third option, namely laying down a list of scenarios where the reliance on a particular legal ground is prohibited or restricted, is the most practical solution in the current legal framework. This is in particular the case with regard to consent when it comes to sensitive data: Article 9(2)(a) has explicitly allowed Union or Member State law to stipulate the purposes for which processing of sensitive data cannot be justified by explicit consent.[32] Under this provision, EU and national policymakers can, in practice, draw up a data protection blacklist without amending the GDPR. As regards non-sensitive personal data, it would be more difficult to work around the legal grounds without changing the law at EU level. As a matter of fact, it has been confirmed in CJEU case-law that the list of legitimising bases for processing of personal data is exhaustive and restrictive, meaning that no legal basis may be added, removed or amended by Member State laws.[33] That said, given the advisory powers conferred to national and EU data protection authorities,[34] they may in practice issue opinions laying down a list of activities that they consider in violation of the principles or rules set out by the GDPR.[35] For

[32] ibid art 9(2)(a).
[33] See Joined cases C-468/10 and C-469/10 *ASNEF* [2011] OJ C 25/18 paras 30-39; Case C-582/14 *Breyer* [2016] OJ C 475/3 paras 56-58.
[34] GDPR, arts 58(3)(b), 70(1)(b) & (e).
[35] For example, with regard to online political advertising, a number of national data protection authorities have already issued guidelines on the lawfulness of such practices. See

example, such a list can be crafted based on the regulator's interpretation of what data uses would count as 'fair' under the fairness principle,[36] or what purposes would count as 'legitimate' under the purpose limitation principle.[37] While such opinions are not legally binding,[38] they would serve as helpful clarification on the regulators' position on the matters that should not be allowed even when consent or other legal bases are in place.

The next question then would inevitably be what should be included in such a blacklist. It would go beyond the scope of this study to attempt to come up with a complete list of the blacklist scenarios. In any case, it would take a lot of further research and public contemplation before such decisions can be made. Restricting individual choice by means of a blacklist requires a compelling justification. Having said that, recalling the potential risks arising from the use of personal data (in Chapter 3) and the flaws of the regulatory model (in Chapter 5), it is still possible to sketch out a rough analytical framework for policymakers to locate the latent harms and to argue the case for excluding individual choice in the form of consent. Following the structure of Chapter 5, such a framework can be presented three-dimensionally with specific focus on autonomous interest, interpersonal (or commercial) interest and political interest.

In terms of autonomous interest, it has been illustrated why sometimes individuals can be better protected by limiting their choice. Recent developments in technologies as well as commercial practices have shown an even more pressing need to introduce a blacklist for certain categories of data processing. For instance, it is already suggested that emotion detection on

---

European Data Protection Supervisor, 'Opinion 3/2018 on online manipulation and personal data' (2018), 6.

[36] GDPR, art 5(1)(a).

[37] ibid art 5(1)(b).

[38] ibid Recital 143.

social media or for advertising purposes should be prohibited.[39] It is believed that, given the power asymmetries in the data subject-controller relationship, 'practical realities belie the libertarian assumption of the capacity and rationality of the data subjects *vis-à-vis* decision-making and also the potential for market equalising effects.'[40] Here, a blacklist approach might be justified on the basis that the activities involved are so sensitive to fundamental interest of the data subject and yet the imbalance of power is so severe that counting on individual choice would not sufficiently address the market failure. For sensitive data that would reveal the most intimate aspects of individuals, it can be argued that it should never be used for, say, advertising purposes, even if explicit consent has been obtained. The recent scandal concerning dating app Grindr sharing HIV status data to third-party advertisers[41] also shows how critical it can be to have such prohibitive measures in place.

As for interpersonal interests, an argument can be made for restrictions on consent for certain categories of processing where the consent may have an impact on other people, which may happen in three slightly different settings. The first one concerns a situation where the data subject's consent is made to a set of data *also* pertaining to others. In fact, this is what happens when a social media user consents to the sharing of their friend list data to a third-party application.[42] Although technically the disclosure of relational data concerning

---

[39] Damian Clifford, 'Citizen-consumers in a Personalised Galaxy: Emotion Influenced Decision-making, a True Path to the Dark Side?' (2017) 31/2017 CiTiP Working Paper <http://ssrn.com/abstract=3037425> accessed 21 February 2018.

[40] ibid 28.

[41] Maya Oppenheim, 'Grindr to Stop Sharing HIV Status of Users with Third-party Companies after Fierce Criticism' *The Independent* (3 April 2018) <https://www.independent.co.uk/life-style/gadgets-and-tech/grindr-hiv-status-users-third-party-companies-stop-sharing-criticism-gay-dating-app-a8286366.html> accessed 3 May 2018.

[42] See Laura Hautala, 'Facebook Privacy Settings Make You Work to Stop the Data Sharing' *CNET* (22 March 2018) <https://www.cnet.com/news/how-to-stop-sharing-facebook-data-after-cambridge-analytica-mess/> accessed 3 May 2018; Josh Constine, 'Facebook Is Shutting

more than one data subject would require consent from *all* concerned parties,[43] this could have been made clearer in a blacklist approach. The second scenario touches upon the lack of choice as highlighted in Chapter 5. In theory, one's consent to a particular service, data processing model and data network would potentially cause psychological and economic effect that would exacerbate the concentration on these dimensions.[44] Such forms of lack of diversity in alternative options would in turn make 'choice' less meaningful in a commercial context. For example, an individual's consent for the sharing of their social media data to an insurer, something that has actually been experimented by an insurance company with discounts offered to pilot users,[45] would increase the popularity of this practice and hence the bargaining power of the insurer. Such voluntary disclosure of personal data, once reaching critical mass, would place the insurer in such an advantageous position that it can practically refuse to provide a quote to those who do not consent, which in turn limits the alternative options available to future customers. In other words, individual consent may end up restricting choice to the collective. A third, and even more subtle account for the interpersonal effects of individual consent is manifest in the case where one's disclosure of their own information may contribute to the data controller's better knowledge of the entire group,[46] and therefore result in negative consequences on the members of the group.

Down Its API for Giving Your Friends' Data to Apps' *TechCrunch* (28 April 2015) <https://techcrunch.com/2015/04/28/facebook-api-shut-down/> accessed 3 May 2018.

[43] See Data Protection Act 1998, s 7(4)-(6); *Durant v FSA* [2003] EWCA Civ 1746 paras 60-67; *DB v GMC* [2016] EWHC 2331 (QB) para 88; Information Commissioner's Office, 'Determining What Is Personal Data' (2012), 24-25.

[44] See Mark MacCarthy, 'New Directions in Privacy: Disclosure, Unfairness and Externalities' (2011) 6(3) I/S: A Journal of Law and Policy 425, 459-461.

[45] Graham Ruddick, 'Facebook Forces Admiral to Pull Plan to Price Car Insurance Based on Posts' *The Guardian* (2 November 2016) <https://www.theguardian.com/money/2016/nov/02/facebook-admiral-car-insurance-privacy-data> accessed 30 May 2018.

[46] Yoan Hermstrüwer, 'Contracting Around Privacy: The (Behavioral) Law and Economics of Consent and Big Data' (2017) 7(3) JIPITEC 9, 12.

For example, sharing one's own behavioural data as well as their shopping preferences to an advertiser may have a collective impact on the group of people sharing similar behavioural pattern with them. This has been explored by some researchers with the concept of 'group privacy'[47] or 'collective privacy'.[48]

The last possible justification for the blacklist approach concerns political interests. As underlined in Chapters 3 and 5, the use of personal data in an OBA context not just concerns individual or commercial interests, but can also involve political considerations. If the personal choice of a data subject — who is also a citizen — proves to contribute to a process that would jeopardise the societal or public good, then it is at least arguable that such choice should be restricted. It is for this reason that in the case of online tracking, an argument has been made that at least in some scenarios, the practice of making provision of service conditional on the consent to tracking should be prohibited.[49] This argument can be extended to a wider context of the use of personal data for political purposes. The revelations of Cambridge Analytica show that there should be further restrictions on how personal data can be used for political campaigning. Compared to how common it is for states to impose complex regulatory regimes on how campaigners may use funding,[50] the legal

[47] Lanah Kammourieh and others, 'Group Privacy in the Age of Big Data' in Linnet Taylor, Luciano Floridi and Bart Van der Sloot (eds), *Group Privacy: New Challenges of Data Technologies* (Springer 2017).
[48] Alessandro Mantelero, 'From Group Privacy to Collective Privacy: Towards a New Dimension of Privacy and Data Protection in the Big Data Era' in Linnet Taylor, Luciano Floridi and Bart Van der Sloot (eds), *Group Privacy: New Challenges of Data Technologies* (Springer 2017).
[49] Frederik J Zuiderveen Borgesius and others, 'Tracking Walls, Take-It-Or-Leave-It Choices, the GDPR, and the ePrivacy Regulation' (2017) 3(3) European Data Protection Law Review 353, 364.
[50] European Parliament, *Party Financing and Referendum Campaigns in EU Member States* (2015) <http://www.europarl.europa.eu/RegData/etudes/STUD/2015/519217/IPOL_STU(2015)519217_EN.pdf> accessed 18 April 2018.

regulation on how they may use data is much less developed. The focus is usually on sharing of voter lists and the obligations are usually imposed on political parties.[51] This means, for example, the use of behavioural data on social media by a third-party, as in the case of Cambridge Analytica, is still largely beyond the reach of traditional control over campaigning use of personal data. Therefore, the option of expressly blacklisting at least some of these uses should be kept open to future debates.

The reflections on autonomous, interpersonal and political interests above are meant to provide a general framework for policymakers to consider the necessity and feasibility of adopting certain blacklist measures, although some of these theories might need to be further developed in future research. Again, it is not this study's task to come up with an exhaustive blacklist of all types of data processing that should be prohibited. What is intended is merely to make it clear that the blacklist approach should remain on the table in future discussions and there can be good reasons to advocate this approach in specific contexts.

(c) Limiting the expanding: Constraints on the expansion of white- and blacklist

Both the white- and blacklist approaches are characterised by the fact that by means of setting out mandatory rules, the law already precludes individual choice on certain matters regarding the use of personal data. The theoretic connection between autonomy and choice when it comes to personal data forms an essential part in Chapter 5. It is concluded that, while personal decisions on whether and how one's personal data may be used for a specific purpose play an important part in the actualisation of autonomy, such

[51] Colin J. Bennett, 'Voter Databases, Micro-targeting, and Data Protection Law: Can Political Parties Campaign in Europe as They Do in North America?' (2016) 6(4) International Data Privacy Law, 265-267.

decisions (in particular in the form of consent) are not the only way, and sometimes not even the best way, to achieve genuine autonomy. Collective decisions in the form of compulsory legal requirements may under some circumstances legitimately override individual choice on these matters. It is not uncommon for a social preference to override the individual preference on certain occassions. As regards white- or blacklist measures, it has been shown above that there can be justifications for the adoption of such social preferences. Those measures are defensible not just in the collective or public interest, but sometimes also with the aim to improve protection of individual interests.

However, it should be borne in mind that, apart from consent, drawing up such a white- or blacklist represents a strong form of intervention of law on matters that are arguably better decided by the collective than by individuals. To the extent that data subjects are denied the possibility to exercise control over the fate of their data, the blacklist and whitelist approaches would create a strong sense of heavy-handed, paternalistic interventionism, which requires a compelling case for the adoption of such measures. In other words, policymakers must clearly articulate why leaving such decisions to individuals is likely to lead to some sort of failure that would cause harms to the individuals themselves, a group of individuals or public policy interests, either directly or indirectly. In addition, any decision to include a specific category of data use on the white- or blacklist should be subject to the general constitutional constraints so as to ensure the power of laying down such lists is not abused by the policymakers or regulators.

One such constitutional constraint — indeed the most important one — derives from fundamental rights guaranteed by the Charter. As with any acts by EU institutions, a legislative decision to authorise or prohibit certain types

of data use would be invalid if it is found in breach of the Charter.[52] Since data protection has been officially recognised by the Charter as a fundamental right, the protection of and the restriction on this right are both subject to the general legal framework for fundamental rights. It has been repeatedly emphasised by the CJEU that the measures adopted by the EU concerning processing of personal data must be proportionate to the legitimate aim pursued.[53] The Court has consistently invoked Article 52(1) of the Charter to stress that limitations on fundamental rights may be allowed if and only if, '[s]ubject to the principle of proportionality, they are necessary and genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of others.'[54] It follows that, for the scrutiny of a specific measure, policymakers must determine whether the interest they claim (which could be one of those mentioned in the previous section) is balanced against the fundamental rights of the data subject or any third-parties.[55] The principle of proportionality, in short, requires that the measures sought: a) 'be appropriate for attaining the legitimate objectives'; and b) 'do not exceed the limits of what is appropriate and necessary'.[56] Such an assessment would of course entail balancing against a variety of interests and rights. In fact, the GDPR itself has spelt out a series of the fundamental rights that should be taken into consideration, including, in particular and *inter alia*, 'freedom of thought, conscience and religion, freedom of expression and information, [and]

[52] See Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland and Seitlinger and Others* [2014] OJ C 175/6; Case C-362/14 *Schrems* [2015] OJ C 398/5.
[53] Schecke, paras 65-89; Digital Rights Ireland, 38-71; Tele2, 94-104.Joined Cases C-92/09 and C-93/09 *Volker und Markus Schecke and Eifert* [2010] OJ C 13/6 paras 65-89; *Breyer* (n 33) paras 38-71; Joined Cases C-203/15 and C-698/15 *Tele2 Sverige* [2016] OJ C 53/11 paras 94-104.
[54] Charter of Fundamental Rights of the European Union [2012] OJ C326/391 ('Charter'), art 52(1).
[55] *Schecke* (n 53) para 77.
[56] Digital Rights Ireland , para 46.*Digital Rights Ireland* (n 52) para 46.

freedom to conduct a business'.[57] If after a thorough evaluation it turns out that the measures are either inappropriate or unnecessary, policymakers should consider alternative approaches instead, some of which will be discussed in the next section.

6.2 Beyond black and white: Diversification of the regulatory toolbox

6.2.1 Lessons from behavioural economics and psychology

While the previous section has covered most of the safeguards on the data protection white- and blacklist — both those already included in the GDPR and those potentially on an expanded list — one special category of measures should be discussed in more detail: consent. Consent deserves further analysis not just because it is less paternalistic in nature than other black- or whitelist measures, but also because it is the most commonly-used (and thus the most controversial) legal basis. In fact, as concluded in Chapter 4, the GDPR is essentially built on the regulatory model of 'consent + necessity 2.0'. Voluntary mechanisms, such as consent, are actually more common throughout the GDPR as the means by which data processing is authorised. One of the reasons why consent is a popular regulatory option may be that it is closer to the idea of personal choice, and personal choice is conceptually connected to the idea of individual autonomy. However, most of the last chapter has been devoted to the discussion of why consent would not be an effective regulatory tool in the context of OBA: It is rigid and flat, and more importantly, it fails to diversify choice. The way choices are made available and framed matter significantly to individual decision-making in that it imposes not just the constraints on available options, but also the mental modality of choosing

[57] GDPR, Recital 4.

between those options. This idea is captured by behavioural economists with the terminology of 'choice architecture'.

One of the most-cited discussions on choice architecture is the book *Nudge* by Thaler and Sunstein. In their book, they argue: 'If you indirectly influence the choices other people make, you are a choice architect. And since the choices you are influencing are going to be made by Humans, you will want your architecture to reflect a good understanding of how humans behave.' A classic demonstration of how choice architecture influences human behaviour is defaults, which makes a perfect point in the case of regulatory approaches for data protection policymakers. The entire debate surrounding the interpretation between 'opt-in' and 'opt-out' consent under the ePrivacy Directive is in essence all about setting a default. The default position of a choice architecture would have an effective impact on the outcome of individual choice, even when the options are the same under various architectures.

Behavioural psychologists have offered a great deal of theoretical accounts for such phenomena. In his Nobel Prize-wining study of bounded rationality, Kahneman summarises the rules governing intuitive choices.[58] A striking conclusion is that individual choices consistently suffer from systematic biases.[59] Based on such ground-breaking works, further studies have been conducted to prove both theoretically and empirically how human decision-making is subject to various forms of external influences, including choice architecture. Sunstein, for example, theorises such biases by outlining the situations where individual choices do not promote their own ends, or in his terms, where there are behavioural market failures: a) present bias and time

[58] Daniel Kahneman, 'Maps of Bounded Rationality: Psychology for Behavioral Economics' (2003) 93(5) The American Economic Review 1449.
[59] ibid.

inconsistency; b) ignoring shrouded attributes; c) unrealistic optimism; and d) misestimating probability.[60] Translated into the context of choice over uses of personal data, these theories may further explain the failure of consent as pointed out in the last chapter.

However, it should be noted that the theory of choice architecture is subject to quite some scepticism.[61] It is criticised both for being too tough and for being too soft. On the one hand, the legitimacy of regulation by means of adopting a particular choice architecture is challenged for its potential interference with autonomy. It is considered at best a pretentious or untransparent version of paternalism,[62] and at worse a blatant manipulation on individual autonomy.[63] On the other hand, choice architecture is also believed to be less effective than mandates in protecting individuals from irrational decisions.[64] Certain responses are offered by advocates of the theory, including, *inter alia*, that while choice architecture can be manipulative where the architect is malicious, in certain areas it can promote autonomy and/or welfare.[65]

---

[60] Cass R. Sunstein, 'The Storrs Lectures: Behavioral Economics and Paternalism' (2013) 122 The Yale Law Journal 1826, 1842-1852. See also Frederik Zuiderveen Borgesius, 'Behavioural Sciences and the Regulation of Privacy on the Internet' in Alberto Alemanno and Anne-Lise Sibony (eds), *Nudge and the Law: A European Perspective* (Hart Publishing 2015) 197-198.

[61] Peter John, *How Far to Nudge? Assessing Behavioural Public Policy* (Edward Elgar 2018) 88-107.

[62] Uwe Volkmann, 'Nudging, Education, Paternalism: A Philosophical Perspective from the Old Europe' in Alexandra Kemmerer and others (eds), *Choice Architecture in Democracies : Exploring the Legitimacy of Nudging* (Nomos 2016).

[63] Daniel M. Hausman and Brynn Welch, 'Debate: To Nudge or Not to Nudge' (2010) 18(1) The Journal of Political Philosophy; Hans Michael Heinig, 'Autonomy vs. Technocracy: Libertarian Paternalism Revisited' in Alexandra Kemmerer and others (eds), *Choice Architecture in Democracies : Exploring the Legitimacy of Nudging* (Nomos 2016); Christopher McCrudden and Jeff King, 'The Dark Side of Nudging: The Ethics, Political Economy, and Law of Libertarian Paternalism' in Alexandra Kemmerer and others (eds), *Choice Architecture in Democracies : Exploring the Legitimacy of Nudging* (Nomos 2016).

[64] McCrudden and King (n 63).

[65] Cass R. Sunstein, *The Ethics of Influence: Government in the Age of Behavioral Science* (Cambridge University Press 2016) 84-87; Cass R. Sunstein, 'The Ethics of Choice Architecture' in Alexandra Kemmerer and others (eds), *Choice Architecture in Democracies : Exploring the Legitimacy of Nudging* (Nomos 2016).

It should be made clear from the outset that this study does not advocate the notion of choice architecture — or its practical implementation, which includes the greylist approaches as will be discussed below — *over* or *as a substitute for* traditional ways of regulation. In fact, many opponents of the choice architecture theory do not object to the complementary potential of this approach;[66] rather, what they oppose is the proposition of abandoning the mandatory rules.[67] In this respect, the regulatory toolbox envisaged by this study does not have to be subject to these criticisms, because, for one thing, choice architecture is considered merely *one of* the options that is not necessarily superior to others, and for another, mandatory approaches, such as the white- and blacklist, are in fact equally supported by the findings of the previous section. More importantly, the promotion of a particular choice architecture model does not mean that a choice architecture model would not exist without such an approach. When regulation is absent, the choice architecture would be decided by private entities,[68] and in some cases, choice architecture may even exist without an architect.[69] Taking consent as an example, when data protection law does not specify whether consent should be 'opt-in' or 'opt-out', it would be up to the data controller to decide which solution to adopt; but in either case, the data subject would face a choice architecture and their decision may be conditioned accordingly.

It is for this reason that the intention of this study is neither to dive into the rich literature of behavioural economics and psychology, nor to pick sides in

---

[66] John (n 61) 106-107,121.

[67] Sabino Cassese, 'Exploring the Legitimacy of Nudging' in Alexandra Kemmerer and others (eds), *Choice Architecture in Democracies : Exploring the Legitimacy of Nudging* (Nomos 2016); McCrudden and King (n 63).

[68] Woodrow Hartzog, *Privacy's Blueprint: The Battle to Control the Design of New Technologies* (Harvard University Press 2018) 52.

[69] See Robert Nozick, 'Invisible-Hand Explanations' (1994) 84(2) The American Economic Review 314.

the debate over the superiority of the regulatory approaches based on the choice architecture theory. What is intended is simply to present such a possibility as an option among others, which has both its strengths but also limitations. Keeping this regulatory instrument in the toolbox does not mean that other instruments should be discarded or less-favoured. Depending on the social consensus and the judgment of the policymakers, the role of this approach can be merely additional or supportive to the existing authorisation-consent-prohibition trichotomy. From such an open-ended perspective, the lessons from behavioural economics and psychology might sound less ambitious: Human choice is subject to behavioural rules and vulnerable to 'architectural manipulation' by others, which may in practice counteract the assumption of rationality on which legislative and regulatory activity is based. Regulators should thus recognise such rules and, where necessary, either make use of them or defend individuals' decision-making capabilities from undue influences. Leaving these opportunities unchecked and solely exploited by private entities would unfairly tip the balance in favour of those entities.[70] As Sunstein puts it, 'choice architecture is inevitable and […] behavioral market failures do, in fact, justify certain forms of paternalism.'[71] However, given the risks of manipulation even by public bodies, it is a common view that the use of such measures must be subject to strict substantive and procedural checks and balances, such as the proportionality test.[72]

---

[70] Ian Brown and Christopher T. Marsden, *Regulating Code : Good Governance and Better Regulation in the Information Age* (The MIT Press 2013) 31-32.

[71] Cass R. Sunstein, *Why Nudge? The Politics of Libertarian Paternalism* (Yale University Press 2014) 16.

[72] Anne van Aaken, 'Constitutional Limits to Paternalistic Nudging: A Proportionality Assessment' in Alexandra Kemmerer and others (eds), *Choice Architecture in Democracies : Exploring the Legitimacy of Nudging* (Nomos 2016); Gertrude Lübbe-Wolff, 'Constitutional Limits to Health-Related Nudging – a Matter of Balancing' in Alexandra Kemmerer and others (eds), *Choice Architecture in Democracies : Exploring the Legitimacy of Nudging* (Nomos 2016).

6.2.2 Nudging and intervening: Libertarian paternalism and more

(a) Nudges

Paternalism has a bad reputation. It is sometimes associated with coercion[73] and hence the impression that individual choice is disrespected and even deprived. Such association is unnecessary and unhelpful,[74] at least in *some* variations of paternalism. Libertarian paternalism, advanced by Thaler and Sunstein, is said to be one such variation that, on the one hand, promotes the idea that people should be free to choose, and on the other, argues for institutional efforts to steer their choices towards improvement of their lives.[75] Libertarian paternalism is essentially one of the approaches 'that affect choices without coercion'.[76] The way libertarian paternalism achieves the seemingly contradictory dual objectives is leaving the final decision open to individuals but at the same time tweaking around the choice architecture.

In terms of the authorisation scheme for a data protection regime, the design of choice architecture may mean making decisions on a combination of a variety of influence approaches. The most straightforward pair of options are, of course, the so-called 'opt-in' and 'opt-out' consent. A general guiding principle for policymakers to decide on these approaches would be picking those rules that 'reflect the likely choices of informed people'[77] with a rational level of self and collective interest. One of the keywords here is 'informed', meaning that, considering the common cognitive fallacies that would impede individual choice, policymakers should exercise their judgment on what a

[73] N. Fotion, 'Paternalism' (1979) 89(2) Ethics 191, 195.
[74] Julie E. Cohen, 'Turning Privacy Inside Out' (2018) 17 <https://ssrn.com/abstract=3162178> accessed 16 May 2018.
[75] Richard H. Thaler and Cass R. Sunstein, *Nudge: Improving Decisions About Health, Wealth, and Happiness* (Yale University Press 2008) 5.
[76] Sunstein, *Why Nudge?* (n 71) 20.
[77] ibid 138.

rational, informed person would most likely choose, instead of simply following the actual or predicted behaviour of the majority. Such a judgment may lead to a conclusion in favour of 'opt-in' consent, or one in favour of 'opt-out' consent. Where a reasonable conclusion can be safely reached and it is technically possible to treat different scenarios differently, it would arguably be justifiable for the policymakers to assign the proper approach to the given set of scenarios.

It should be noted that the binary of 'opt-in' or 'opt-out' in the ongoing policy debates may have caused an impediment to governmental or industrial initiatives to develop a third path. The highly limited achievement, if any, of the browser-based solutions to use of cookies, as discussed in Chapter 4, has shown how potentially helpful innovations may be restrained due to the entrenchment of the policy choice between the 'opt-in' and 'opt-out' approaches. The development of technical standards for default browser settings or the 'Do Not Track' feature is largely seen by the industry as a technical solution to simplify the way meaningful consent is obtained, whereas European regulators are deeply concerned about the risk of exploitation by the industry to collect massive amounts of data. However, the unmet expectations are not necessarily irreconcilable. If the industry and regulators can come to an agreement that for certain types of low-risk cookie uses, a relaxed, streamlined regime may be allowed even though it might not fully satisfy the arguably harsh 'opt-in' consent requirement, then there would be a considerable scope where the gap of expectations can be bridged. For example, it can be argued that since first-party cookies without storing sensitive data are less dangerous, a browser-based implementation of opt-out consent may be allowed, which would not only free Internet users from a great deal of cookie consent, but also create incentives for service providers to

consider using first-party cookies only. However, for higher-risk uses of cookies, such as, say, third-party cookies for advertising purposes, it might be sensible to adhere to the regulators' more restrictive interpretation of the conditions for valid consent. Nevertheless, it should be clear that applying either 'opt-in' or 'opt-out' consent to all matters will not be a helpful approach.

Apart from default rules, libertarian paternalism may take other forms in the regulatory toolbox. In fact, when explaining the potential diversity of nudges, Sunstein envisages a list of 13 strategies employable by a hypothetic cigarette regulator.[78] For those that qualify as soft paternalist measures, the approaches mainly involve provision of information, displaying warnings and making purchase more difficult.[79] All these measures can be transferred to a context of use of personal data for online advertising, or generally, commercial use of big data.

For provision of information, the discussion may go all the way back to the theoretical significance and practical failure of the transparency principle. Mandatory disclosure of certain categories of information is essential in that it underlies informed decision-making. However, such transparency requirements are often criticised for their insufficiency in truly informing people's choice.[80] From a regulatory point of view, however, information does not always have to be truly informing, nor does it always have to be in place.[81] It can be an option for regulators to leverage. For instance, it is argued that

[78] Sunstein, 'The Storrs Lectures' (n 60) 1864-1865.
[79] ibid.
[80] See Omri Ben-Shahar and Carl E. Schneider, 'The Failure of Mandated Disclosure' (2011) 159 University of Pennsylvania Law Review 647; Lilian Edwards and Michael Veale, 'Slave to the Algorithm? Why a "Right to an Explanation" Is Probably Not the Remedy You Are Looking For' (2017) 16(1) Duke Law and Technology Review 1.
[81] In an experiment-based study, it is even proved that salient information and prior consent may create the unexpected chilling effect. See Yoan Hermstrüwera and Stephan Dickert, 'Sharing is Daring: An Experiment on Consent, Chilling Effects and a Salient Privacy Nudge' (2017) 51 International Review of Law and Economics 38.

when making privacy-related decisions, individuals do not need to be fully informed; sometimes being sceptical is enough for discouraging certain behaviour.[82] For this reason, disclosure of information can be seen as a tool, a regulatory tool designable by policymakers. They can design what information and how it should be conveyed to data subjects. One promising solution is standardised icons,[83] which should function in a way similar to nutrition facts labels that give consumers a straightforward overview of the information they need most.

Warnings are a similar category of nudges, except that it works by producing psychic costs to individuals.[84] Such costs can be important signals of the potential risks of data processing concerned, which, as illustrated in Chapter 5, forms a dispensable part of the role of data protection law. Therefore, the quality and quantity of such warnings are crucial[85] in that proper implementation of this approach would enable data subjects to tell the important decisions from the trivial ones, so as to avoid the 'warning fatigue'. Making consent harder to give would generate another form of transaction costs, which, again, can be a regulatory possibility that nudges individual behaviour towards a certain pattern or sends out signals of the importance of the decision being made. This may include, for example, a two-step confirmation or a compulsory periodic review of consent. The combination of all these regulatory options would allow policymakers or regulators to create a range of choice architecture applicable to different circumstances.

[82] Hartzog (n 68) 176.
[83] GDPR, art 12(7).
[84] Sunstein, *Why Nudge?* (n 71) 57.
[85] Hartzog (n 68) 129.

(b) Interventions

While libertarian paternalism aims to nudge individuals into choosing the most informed option, sometimes neither of the options is close to an optimal solution. Choice architecture may effectively influence how people decide between the options offered, but it does not address the issues arising from the lack of the ideal option. The failure of data protection law in diversifying choice has been made clear in Chapter 5 in particular when the effectiveness of consent is examined in terms of protecting data subjects as consumers. If such diversity of choice proves to be crucial in protecting the interpersonal (commercial) interest of data subjects, regulators may consider adopting stronger forms of paternalism, including interventions in the scope of selections for key decisions.

One potential instrument of this kind is to make alternative business models available to data subjects. This can be achieved through a hard version by, for instance, making it compulsory for service providers to offer a paid-for option if the user prefers not to have their data shared for advertising purposes. A soft version would be providing that the availability of the paid-for option may, in the course of the 'freely given' test, demonstrate the non-conditionality between the access to main services and the consent to use of personal data. Extra precaution is however needed for this approach to avoid the potential side-effect of creating the privacy-haves/-have-nots divide. As flagged up in the previous chapter, policymakers will need to decide carefully how not to compromise fundamental interests of data subjects.

Another instrument in the toolbox is to make the option of opting out of certain types of data processing always available and accessible to data subjects. For instance, data protection or competition law may require that any merger of user accounts across services can be allowed only if the user has given additional consent and only if the refusal to consent would not affect

their continued use of either service.[86] Policymakers may also decide that any material improvement in certain profiling techniques, even if compatible with the original purpose, must be justified by consent, the refusal of which, again, should not impact the access to the service. A dashboard-like control panel might be also helpful for users to keep part of the data use off or separate. In fact, these regulatory tools are more or less available if certain elements the GDPR are properly applied together, such as purpose limitation and the 'specific' and 'freely given' requirements of consent, except that these mechanisms are more general and do not target specific contexts. As will be discussed below, list-based approaches with specific scenarios in mind are sometimes better policy choice than adhering only to the arguably blunt, one-size-fits-all principle-based approaches.

The discussion here will not go further in search of all possible forms of nudges and interventions, nor will I attempt to explore the strengths and weaknesses of each instrument to find out where they are best positioned to address particular circumstances. These tasks will have to be left to future research. The key point here, however, should be clear enough that there are more delicate, subtle tools in the data protection toolbox that can be useful for dealing with big data issues but are not fully realised by the GDPR.

(c) Nudges and interventions as technology and market mechanisms

For data protection policymakers, various forms of nudges and interventions are not isolated regulatory instruments. Instead, those specific measures should be mapped out within the overarching policy framework. This would allow policymakers not just to identify potentially useful approaches, but also to compare their strengths and weaknesses against alternative options. A helpful analytical framework where nudges and interventions can be located

[86] Cohen (n 74) 6.

would involve a holistic view of the interactions between different regulatory paradigms. The relationships between the parallel yet interrelated regulatory paradigms of technology, market and law have been discussed in great depth in Lessig's book *Code*. He explains how law may affect behaviour by either regulating directly or indirectly through regulating technology (or, in his terms, architecture) and market.[87]

To illustrate how indirect regulation through technology works, the promotion of the use of seatbelts is given as an example. If this is an objective pursued by a government, the government may regulate citizens' behaviour directly by passing a law to require the wearing of seatbelts, or it may do it indirectly by regulating technology, such as mandating automatic seatbelts.[88] In the case of regulating online privacy, however, this is slightly different. The subjects of the regulation — namely the data controllers — are also the designers of an architecture that would have nudge effects on others — data subjects. As such, when discussing direct and indirect regulation in this context, it should be kept in mind that if it is the data subjects' behaviour that is meant to be regulated, there are actually three entry points where legal regulation may exert influence. First, law may directly dictate that data subjects make a certain choice (through a whitelist or blacklist). Second, law may indirectly dictate that data controllers adopt a certain choice architecture that would influence individual choice (through imposing consent or necessity requirement). Third, and even more indirectly, law may impose a spectrum of choice architectures with different legal consequences that would influence the *data controller's choice* of choice architecture (through offering a range of greylists, as will be discussed below).

---

[87] Lawrence Lessig, *Code: Version 2.0* (2nd edn, Basic Books 2006) 127-130.
[88] ibid 130.

The distinction between the second and third layers would become even more evident when they are respectively considered conduct-based and design-based regulation, a dichotomy advanced by Hartzog when calling for diversifying privacy law's approach.[89] Whereas regulating data controllers' conduct is the traditional approach of data protection law, regulating the design/architecture that *they create* and the design/architecture that *affects their behaviour* may be a new approach that policymakers should arguably embrace, or at least consider. If data controllers are in effect choice architects, then it would make much sense for policymakers to act as an architect of architects, or *meta-architect*. By acting as meta-architects, policymakers would effectively create a two-layer market mechanism. On the direct level, policymakers create market incentives for data controllers to adopt choice architectures that are more privacy-friendly. On the indirect level, those choice architectures in turn create positive or negative incentives to individual decisions on various types of data uses.

For the incentives to work, there needs to be a degree of uncertainty on the scope of each choice architecture. As Hartzog notes, '[a]ny privacy design agenda should adequately ensure the freedom for companies to take reasonable risks, learn from mistakes, and account for other regulatory pressures like social norms and market forces.'[90] The same goes for data controllers' decisions on choice architectures. A range of architectures with varying benefits but also varying price tags are provided to data controllers.[91] If they want a more relaxed authorisation approach, they would need to make sure their practices fall within a lighter-coloured list. They would make their decisions based on how much they value the benefits and costs of the use of

[89] Hartzog (n 68) 82-84.
[90] ibid 85-86.
[91] Brown and Marsden (n 70) 165.

data, and also on their own or others' mistakes — regulators send out signals of the scope of a particular architecture by enforcing the authorisation scheme. The binary of consent and necessity offers a much narrower selection and thus will not achieve the same level of optimisation. The failure to effectively enforce the cookie rule is a case in point: When website operators want to use low-risk cookies to monitor web traffic, they simply cannot do it in a fully lawful way as few users will give opt-in consent. Therefore, even if they are aware that this might be legally questionable, they would probably do it anyway, and they would probably do it for other purposes as well, since it would be unlawful anyway. That is why such propositions as 'treating all data as sensitive data' would be unlikely to be helpful. In order for the market mechanisms to kick in and drive commercial practices towards a more desirable direction, a reasonable range of choices would be indispensable.

Of course, a counter-argument can always be that this might as well lead to a slippery slope for online marketers to exploit the choice architectures. In fact, libertarian paternalists have anticipated the same challenge and they have given their own response.[92] Drawing on some of their key ideas, I try to offer a defence along similar lines. First, the potential benefits of this approach as illustrated above cannot simply be dismissed by the fear of a hypothetical slippery slope. Second, the inclusion of alternative choice architectures (other than 'consent' as currently defined by the GDPR) has not changed the fundamental nature of the data protection regime underpinned by the prohibition-consent-permission trichotomy; it simply further supports it with additional choice architectures. Third, refraining from regulating choice architectures is pointless, as '[a] regulator's choice to address design or to ignore it is also a kind of default, which comes with its own set of

[92] Thaler and Sunstein (n 75) 236-238.

consequences.'[93] Some of the points made here may also find further support in the discussion about the list-based approach in the next section.

6.3 Fifty shades of grey(lists)? Towards a strong list-based approach

6.3.1 Greylists and the list-based approach

All the paternalistic approaches discussed in the previous section, whether libertarian or interventionist, are based on the common ground that individual data subjects have the final say on such decisions. As such, these approaches are distinguishable from other white- or blacklist measures in that the decision of the data subject plays a key role in ascertaining whether a given use of personal data can be allowed. Under the current regulatory model, any data uses that are not blacklisted can be actually justified by consent. To such an extent, consent can be seen as a fallback (or 'catch-all') measure that does not need to come with a list of activities. However, with the nudges and interventions highlighted above, it is conceivable that policymakers may create multiple forms of consent, each with a particular choice architecture possibly featuring various combinations of such nudges and interventions. A simple example would be a data protection law that provides both 'opt-in' and 'opt-out' consent as authorisation mechanisms for different categories of data uses.

That would require data protection law to lay down the scenarios in which (or the conditions under which) a particular form of consent may be used as a valid legal basis. If policymakers put together a list of the activities on which individuals may consent in a specific way, such a list can be considered a greylist — neither completely white (permissive) nor completely black

[93] Hartzog (n 68) 54.

(prohibitive). For the matters included on a greylist, data subjects are allowed to act as their own gatekeepers in the authorisation scheme although various mechanisms might exert their influence one way or another. Since it is possible to diversify data protection greylists with different optional combinations of nudges and interventions, paternalistic approaches should be understood as a continuum with gradual shades of darkness without sharp distinction between similar strategies.[94]

The beginning of this chapter has picked up on the idea of a list-based approach to data protection law, characterised by the codification of groups of data uses into a number of lists representing various mechanisms within the authorisation scheme. Here, a further step is taken towards a strong version of the list-based approach, such that apart from a whitelist and a blacklist, the authorisation scheme also features *a number of* greylists. The theoretical justification for the adoption of a strong list-based approach will be given later in this section. For now, the discussion will be first focused on how policymakers may build up such greylists with the nudges and interventions explored above as the building blocks. In order to bring together the technological, market and legal tools and assign them to the greylists, it would in theory entail a three-step process as illustrated below.

First, policymakers need to locate the available regulatory instruments in the full spectrum of their toolbox. As shown below in Figure 6.1, these instruments — some of which have been highlighted in the previous section — may include both nudges and interventions, and may exert influence either through code or through market. It should be made clear here that the measures included in the figure are merely demonstrative and are not meant to be exhaustive.

[94] Sunstein, *Why Nudge?* (n 71) 56-57.

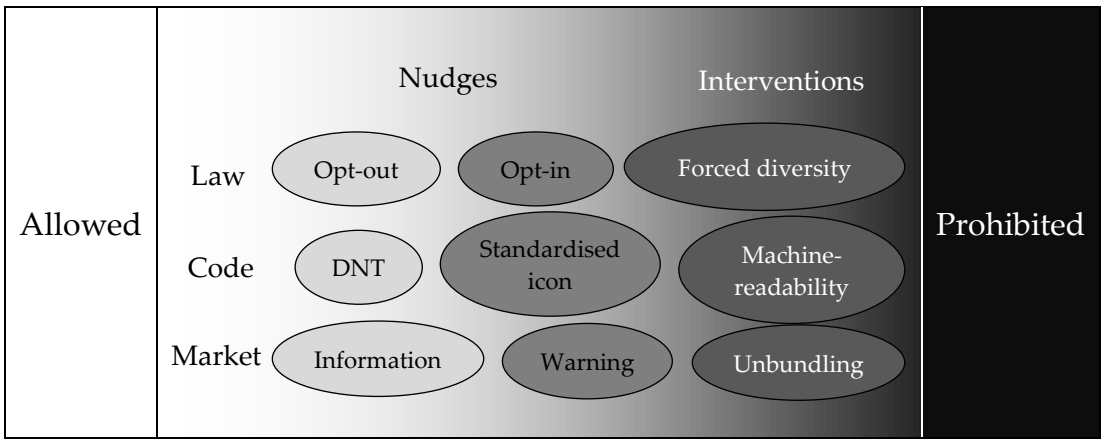

Figure 6.1 Potential policy instruments in the regulatory toolbox

At this point, policymakers may be tempted to leave all these measures available to data controllers, who may determine, on a case-by-case basis, what would be the appropriate combination of these measures for a particular type of data processing, if they seek to rely on consent. Depending on the nature of the activities concerned, the processing in question might fall on a particular point within the spectrum of the shades of grey. With *ex post* oversight by regulators or courts, this would arguably have the advantage of keeping all the options open not just to data controllers but also regulators, who can in turn react to new technologies that would call for a particular form of choice architecture. However, in practice, this would also create a number of issues, such as the potential exploitation by the controller, the over-complexity of the system, the lack of legal certainty, and the burden on regulators and data controllers to decide on the appropriate option. For these reasons, a strong version of the list-based approach, which simplifies the system by reconstructing the spectrum of the grey area into a number of greylists, might be a better regulatory solution. This might suffer from the lower level of flexibility and make the regulatory model less future-proof, but these concerns may be mitigated by the proper distribution of the list-drawing powers between policymakers and regulators, such as delegating certain power to regulators to maintain and update the lists.

To turn the shades of grey into the greylists, two more steps will be needed. The next step would involve the possible combinations of some of those instruments that are compatible. An exemplary series of possible combinations are offered in Figure 6.2, which are arranged in a way that reflects the 'darkness' of each approach: The darker the colour of an approach is, the more prohibitive it tends to be. Again, the figure is merely illustrative in the sense that it does not reflect all possible approaches under an authorisation scheme, and the level of paternalism is not necessarily correctly ordered in the figure. The outcome will largely depend on the specific design of each option as well as the configurations of multiple options.

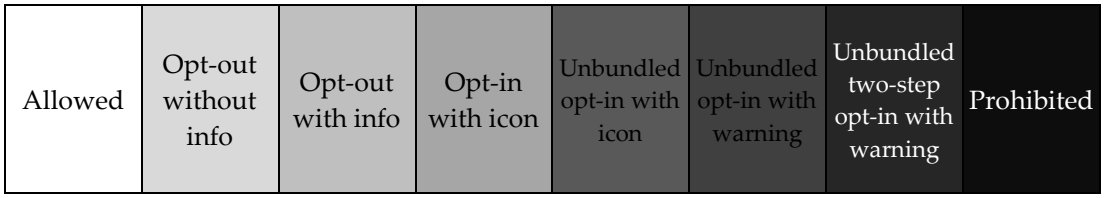

Figure 6.2 List-based approaches arranged in a gradual spectrum

The spectrum of approaches included in Figure 6.2 are already heavily simplified for demonstrative purpose, but are still very complex. In practice, this would raise the issue of putting too much burden on regulators and data controllers to identify and manage different greylists, who might end up being lost in too many options. Another potential problem would be the lack of a catch-all measure in the event that the use of data in question does not fall within the scope of any of these greylists. To overcome these issues, a last step is needed for policymakers to finalise the greylists. As shown below in Figure 6.3, two further improvements should be made to the lists. For one thing, depending on the desirable number of lists in an authorisation scheme and level of complexity of the scheme, policymakers may need to remove certain approaches or to merge some of them. For another, policymakers would also

need to designate a 'baseline' approach for data processing that is not covered by the white-/blacklist or other greylists.

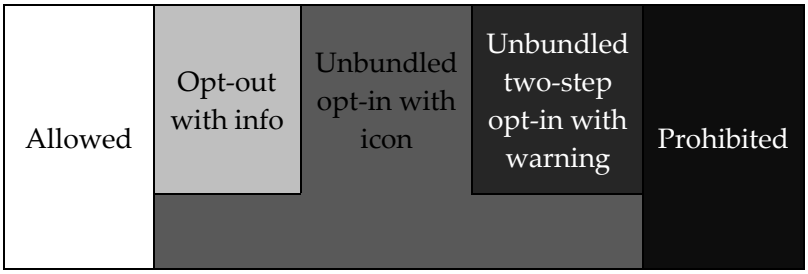

Figure 6.3 A hypothetical set of finalised data protection lists

With these final refinements, it is expected that the authorisation scheme would be less burdensome for both data controllers and data protection regulators. For data controllers, the complexity issue is lessened by the fact that the final scheme is supposed to be streamlined to some extent. Also, since the traditional approaches remain in place in this new scheme, data controllers can always fall back to the 'consent + necessity 2.0' model as it currently stands in the GDPR. Under this new scheme, however, they would be offered a wider range of selection, which, in theory, creates a higher chance for them to choose the most suitable solution. The market will steer the choice of the controllers towards the optimal approach, taking into account the administrative costs and the potential benefits of switching to a different solution. For policymakers, this will indeed increase their burden to scrutinise business practices and decide whether they have chosen the proper authorisation model. However, this will also enable regulators to release more subtle signals on what they consider appropriate safeguards in the given context, for similar data controllers to follow suit. Again, this is helped by the *ex ante* simplification of the authorisation scheme by the policymakers.

Another practical consideration for policymakers would be the choice of an appropriate regulatory instrument to incorporate the greylists into the existing legal framework. This is not the place to expand on all the possible

policy solutions, let alone the pros and cons of each solution. Yet, a number of potential points of entry are briefly offered here for consideration: First, policymakers can of course codify the greylists into the current law, in the form of specific provisions or an annex, or as a new piece of legislation. Second, this can be achieved by granting delegated powers to regulators, who may lay down and maintain such greylists as they see fit.[95] Third, if self-regulation proves effective, regulators may also consider approving the greylists in the form of codes of conduct proposed by representative industrial organisations, which should be legally binding and enforceable on the members of the organisations.[96] Fourth, the greylists can be implemented on a voluntary basis that would create incentives for data controllers to improve the standards of data protection. This approach can possibly be linked to the 'data protection by design' (DPbD) requirement under the GDPR, which requires the data controller to take appropriate technical and organisational measures designed to better comply with the data protection principles and requirements.[97] A greylist in this context may provide guidance on the design of certain nudges, which, when properly implemented, may demonstrate the compliance with the DPbD requirement. The specific conditions should be commensurate with the identified level of risk.

Needless to say, the power of policymakers and regulators to lay down the greylists, just like the white- and blacklist, must be subject to constitutional constraints, including those provided by the human rights framework. On the one hand, a high level of data protection guaranteed by the Charter should not

[95] The GDPR has indeed already empowered the Commission to adopt delegated acts regarding standardised icons and certification mechanisms. See GDPR, arts 12(8), 43(8), 92.
[96] Under the GDPR, associations may prepare codes of conduct and submit them to supervisory authorities for approval. However, while 'the collection of personal data' is one matter that may be covered by the codes of conduct, it does not seem plausible that the consent requirements may deviate from what the GDPR mandates. See ibid arts 40-41.
[97] ibid art 25(1).

be compromised but should always remain the policy objective of introducing such lists. On the other hand, various interests should be weighed and balanced against each other to ensure the chosen regulatory approach is proportionate.

6.3.2 Overcoming 'consent + necessity 2.0' with a strong list-based approach

The strengths of a strong list-based approach are even more evident with the shortcomings of the 'consent + necessity 2.0' regulatory model in mind. With the whitelist, blacklist and several greylists in a strong version of list-based authorisation scheme, the deficiencies of the 'consent + necessity 2.0' model highlighted in the previous chapter can be at least partly alleviated.

To begin with, the rigidity and flatness of consent would be likely to be mitigated by the variety of choice architectures available in the toolbox. At the same time, the failure to diversify choice, especially in terms of diversity of business models and diversity of data networks can be to some extent addressed with the interventions by regulators. Political use of personal data, regardless of the nature of the data concerned, can be placed on the blacklist if that is considered severely detrimental to democracy. The priority and seriousness of the decisions to be made by individuals can also be signified by the appropriate choice of nudges. In the context of OBA, for example, different forms of direct marketing as well as different types of purposes, depending on their nature and risk, may also be treated differently with one of the greylists or blacklist. To sum up, the list-based approach holds great promise to strengthen the existing regulatory framework against the challenges arising from the widespread application of big data. Much more theoretical and empirical work will be needed to prove the extent to which these weaknesses can be overcome with the list-based approach. For now, the task of this study

is to draw attention to such possibilities, which should be open to future debates.

Interest	*Consent*	*Solution*	*Necessity*	*Solution*
Autonomy	Rigidity and flatness	Greylists	Inability to signify priority and seriousness	Nudges
Interpersonal (Commercial)	Failure to diversify choice	Interventions	Inability to capture the uniqueness of the 'direct marketing' purpose	Greylists/ Blacklist
Political	Misplacement of focus on the nature of data	Blacklist	Insufficient categorisation of sensitivity of purpose	Greylists/ Blacklist

Table 6.1: Solutions to shortcomings of the 'consent + necessity 2.0' model

Apart from addressing the existing issues in the current data protection legal framework, the regulatory toolbox also offers another advantage. The flexibility and limited uncertainty it offers will create the helpful tensions between policymakers, regulators, courts and the players in the market. As mentioned above, the way the white-/grey-/blacklists are drawn should allow for a degree of discretion. Such a discretion in effect represents a balanced power distribution in the course of decision-making on what counts as high- or low-risk, and what should be put on which list. Ideally, policymakers are supposed to lay down the whitelist, the blacklist and a number of greylists with particular sets of paternalistic elements. They should also decide on the types of data uses that should be put onto a particular list, which requires some room for interpretation. The regulators will then be able to enforce such lists by examining whether data controllers have chosen the appropriate list for their use of data. Where there is a dispute arising from the enforcement of such lists, a court may step in to decide whether the regulator has exercised their discretion in a proportionate manner. A number of such flexible lists would therefore not just ensure the list-drawing power is not exclusively vested in one decision-maker, but also permits the lists to be slightly stretched to cover

new technologies or new circumstances. Even more importantly, such flexibility creates incentives for technology and market forces to move towards more privacy-friendly direction.

6.3.3 A brief response to potential principle-based criticisms

The white-/grey-/blacklist spectrum may anticipate criticisms from advocates of principle-based approaches, in particular against a strong list-based approach characterised by a variety of greylists. For instance, Koops argues that, '[i]nstead of making data protection law broader and more detailed in how it is to be implemented and enforced, which makes it more complex and more rigid and therewith unrealistic for 21st-century data processing, data protection law should be simplified and focus more on the main underlying principles.'[98] A supporting argument can be that '[t]he complexity of potential consequences and thus the risks are beyond comprehensive regulative instruments such as the law or political decision-making processes.'[99] These arguments seem to oppose a more specific, list-based data protection framework that involves the drawing up of multiple lists. I will provide a number of brief points to explain why this does not have to be the case.

First of all, the distinction between the list- and the principle-based approaches is largely a matter of relativity. Some researchers have classified privacy as a principle-based right whereas data protection is seen as a rule-based right.[100] This means despite the fact that the fundamental principles play a key part in the underlying structure of data protection law, specific rules

---

[98] Bert-Jaap Koops, 'The Trouble with European Data Protection Law' (2014) 4(4) International Data Privacy Law 250, 259.

[99] Leon Hempel and Hans Lammerant, 'Impact Assessments as Negotiated Knowledge' in Serge Gutwirth, Ronald Leenes and Paul De Hert (eds), *Reforming European Data Protection Law* (Springer 2015) 130.

[100] Hielke Hijmans, *The European Union as Guardian of Internet Privacy: The Story of Art 16 TFEU* (Springer 2016) 66.

nevertheless form an indispensable component of the regime. The list-based approach may indeed require more detailed and targeted specifications, but considering its effectiveness in neighbouring areas of law,[101] there is no reason to believe it would not work in data protection law.

Second, the argument for an additional list-based approach does not mean that the principle-based approach is not necessary anymore. In fact, even when the white-/grey-/blacklists have been adopted, the 'safety net' woven by the principles will remain of paramount importance in the data protection framework to provide baseline protection where the situation is not covered by any of the lists.

Third, the proper functioning of a list-based approach is actually dependent on a robust set of the existing principles. The quote above making reference to the complexity of potential consequences and risks has indeed made a strong case, which echoes Hartzog's explanation on why standards are generally better than rules as guidelines for privacy design. The reasons include 'the rapid pace of technological change, the contextual dependency of all privacy problems, and the wealth of common knowledge companies can draw from'.[102] A similar claim can be made here with regard to the list-based approach: Taking into account all these constraints, the effective lists drawn up by policymakers would be those 'articulating goals and erecting boundaries' that reflect 'flexible standards rather than rigid requirements'.[103] Such goals and standards will need to draw inspirations from the values embedded in the data protection principles.

Fourth, both the principle- and the list-based approaches aim to introduce context-specific measures to supplement the principles, rather than to replace

[101] See UCTD; UCPD.
[102] Hartzog (n 68) 122.
[103] ibid 121.

them. The essence of a principle-based data protection law, as Koops argues, is to 'provide a general framework in which the spirit of data protection is clearly visible, in contrast to the EU law's tree-obscuring forest of rules.'[104] This would require 'going back to basics [and] playing other regulatory tunes on different instruments in other legal areas'.[105] Therefore, the fundamental difference between the answers provided by the principle- and the list-based approaches, if any, is that the former attempts to keep data protection purely principle-informed by looking for context-specific measures elsewhere, whereas the latter seeks to incorporate such measures into data protection law.

6.4 Summary: Rebuilding the walls with regulatory tools

In Chapter 3, it is argued that the data protection challenges we are facing today can be captured by the idea of the falling walls of the data world. Now that a potential list-based regulatory model has been outlined, it is time to return to those challenges to see how the proposed solution may possibly help individuals rebuild their own data walls.

The metaphor of walls is a powerful one in that it bridges the gap between our physical experiences of preserving privacy by blocking the observation of others and our digital experiences of protecting data by preventing access by others. A wall represents an architectural barrier that by default prevents information from transferring from one side to the other. Peeking through a wall is not impossible but it would involve extremely high costs. Homes are almost always surrounded by walls, whereas parks rarely have any walls. Glass curtain walls are more common in office building but are much less so for residential buildings. Walls represent people's expected visibility from

[104] Koops (n 98) 259.
[105] ibid 261.

outside in a given space, or in Hartzog's words, peoples' obscurity.[106] Our daily obscurity in the real world is maintained through different layout of walls or other forms of architecture, varying from public space to workplace, and to home. This is why obscurity should be considered a spectrum ranging from completely obscure to completely obvious.[107]

In this context, the list-based approach to data protection is helpful in retaining a similar sort of obscurity in that it does not view the world simply as public and private. By promoting the idea of a diversity of greylists with distinct configurations of nudges and interventions, this approach shares a similar perception with obscurity theories. It also helps rebuilding the walls by making clear demarcation between different lists and giving out clear signals on what one can expect in a given space in terms of the use of their data. This works like when one is going through a door from one room to another, they can clearly see a label on the door specifying what the doors would look like in the next room.

Throughout this chapter, a data protection toolbox has been sketched out, laying down a potentially promising authorisation scheme made up of a number of lists that control the use of personal data. It is suggested that for a data protection framework to function effectively in the age of big data, what we need is perhaps a series of alternative libertarian paternalistic approaches. They are mirrored in a continuum of choice architectures and serve to further strengthen the existing regime. It is believed that through this approach, the shortcomings of the 'consent + necessity 2.0' model can be effectively addressed. Another reason to support this regulatory model is that it can potentially leverage technological and market drivers to improve the

[106] Hartzog (n 68) 110.
[107] ibid 112.

adherence to existing data protection principles. The bottom line is that in a big data world full of subtle influence and complex technologies, policymakers need to give regulators access to a better-equipped data protection toolbox. It remains subject to debates what exactly should be included in the toolbox, but realising such regulatory possibilities and beginning the discussion would be the first step.

Conclusion

It has been a long way here from the very beginning of this study, where the research question was set up, to the development of an improved regulatory toolbox. With six chapters each addressing one particular aspect of the debates surrounding big data and data protection law, the entire inquiry has yielded an answer to the overarching question posed at the outset. Nevertheless, after the in-depth and detailed discussion on such a complex and current issue, it is worth taking one step back to review the logical trajectory following which the investigation has led to the conclusion, as well as to relocate the research findings from the case study of OBA to the bigger picture of big data and data protection law. This conclusion chapter will therefore recapture the main arguments advanced in all previous chapters, and reiterate how such arguments connect to each other and contribute to the main theme of this study. This will be followed by further clarification of the boundaries of the research findings, and their potential transferability in order to inform research and policymaking in other wider contexts. The final section will then acknowledge the limitations of this study, and identify the adjacent fields that are not covered but should be addressed in future research.

Review of research findings

Before summarising the key messages of the main chapters above, it is important that the principal question of inquiry be borne in mind all the way through. The main title of this study — *Data Protection in the Age of Big Data* — should serve as a helpful reminder of the two major keywords: big data and data protection. The two-directional relationship between these two concepts constitutes the main theme of this study: *How should data protection law respond to the challenges arising from the ever-increasing prevalence of big data?* Of course,

given the immensely broad coverage of both concepts, they have each been pinned down to a more specific context. For big data, this means OBA has been employed as a case study throughout all chapters; for data protection, the EU's GDPR has been chosen as the legal framework. Accordingly, the research question can be specified as: *How should data protection law, represented by the EU's General Data Protection Regulation, respond better to the challenges arising from the ever-increasing prevalence of big data in the context of online behavioural advertising?*

In Chapter 1, the enquiry began with a brief introduction to the notion of big data, which, despite the variance in the precise definition, is defined by the high volume, velocity and variety of data. In this regard, OBA indeed represents a helpful instance for the investigation, since the operation of such a system involves the collection, transmission, aggregation, analysis, storage and reuse of massive amounts of data among interconnected data points across devices almost instantaneously. This becomes even more evident as the technical implementation of a typical OBA network was explained within the operational cycle of tracking, profiling and targeting. One example of the sophisticated techniques involved is real-time bidding, which shows the logic of optimisation behind such systems. The complexity and powerfulness of an OBA network is confirmed also from an economic perspective. It has been explained how the few key players in the ecosystem have secured enormous dominance by growing in size, breadth and impact. By means of vertical and horizontal integration, those powerful actors have ensured their involvement in almost all parts of the value chain, which makes their performance better, and in turn their dominance harder to resist.

Such techno-economic realities of the OBA sector provide strong support to certain counter-arguments against some common theories in favour of the

use of big data in an online marketing context. Chapter 2 examined these claims with regard to three separate (though interrelated) aspects: the potential benefits for individuals, the economy and the society. As regards individuals, claims are often made either with reference to people's preference of relevant ads, or to their actual access to free content or services. The former point is usually advocated with survey results suggesting that a large population of Internet users prefer tailored ads, which, as has been shown, is flawed in terms of both methodology and relevance. The latter point concerning the 'free' Internet is less problematic in that free access to online services is indeed of a certain degree of value, although the value has often been exaggerated by industry reports. As for the economy, it is argued that OBA has been making considerable contribution both as an emerging industry on its own, and as part of the bigger digital market. Again, neither of these points can fully justify the industry's position. On the one hand, the development of an industry, whether domestically or internationally, does not automatically override other policy objectives. On the other hand, the statistics cited by industry reports are likely to be overestimated, and have not considered the potential benefits of alternative business models. In terms of the last strand of justifications regarding the potential societal benefits, it is claimed that OBA has been promoting innovation and democracy. The innovation argument is questionable if scrutinised in the light of theories of responsible innovation, which do not consider innovation as something intrinsically desirable. The democracy argument is equally challengeable in that it has largely ignored the threat that OBA may lead to the creation of new forms of bundling. Considering all these aspects, it can be concluded that while OBA, as an application of big data, indeed brings forth certain personal,

economic and social benefits, most of the arguments made by the marketing industry are at best exaggerated and at worst ill-founded.

With some of the potential risks already foreshadowed in the previous chapter, Chapter 3 further identified the fundamental threats that OBA might cause, with a range of theories concerning the interests of individuals and the society. First, while a narrow, classical sense of privacy is often considered violated in the online environment, OBA presents a hard case for the privacy-based risk theory, as the line between private and public space is highly ambiguous and protean. Second, and in a more contemporary vocabulary, informational self-determination is another fundamental value that is considered under threat. Theories of informational self-determination articulate the merits in terms of both individual development of identity and social preservation of pluralism, but in either case, it represents a highly instrumentalist value underpinned by something more fundamental. Third, and as a result, the genuine risks of OBA should be understood with reference to human dignity, which can be discussed in terms of liberty and equality. It is concluded that both values are at stake in the context of OBA as it is likely to create manipulative and discriminatory effects. Fourth, the risk of OBA can be also perceived from an economic perspective, which leads the discussion to the power imbalance between online marketers and Internet users. Such imbalance takes the form of asymmetries in architectural, informational, bargaining and manipulative power. Fifth, from a political point of view, the prevalence of OBA may also undermine the proper functioning of the democratic system by causing chilling effects, creating unfair advantage and generating filter bubbles. To better make sense of how big data has intensified all these risks, a metaphor of 'the falling walls' has been put forward,

explaining the ways in which big data has eliminated the boundaries between various domains of the data world.

An effective data protection framework should be competent to address the risks identified above, but before evaluating the effectiveness of the framework, Chapter 4 provided a doctrinal analysis of some aspects of the current data protection regime that are most closely related to the operation of a typical OBA system. These include the definition of personal data, the legal bases for lawful processing of personal data, the data protection principles, and the restrictions on profiling. The legal analysis in this chapter leads to a number of practical implications: First, the behavioural data collected by OBA systems falls within the definition of personal data, and therefore data protection law applies to related activities. Second, the safest, if not the only, legal ground for OBA-related data uses is consent, which has been further strengthened by the GDPR, but the efficiency will depend on how certain provisions are enforced and how the future ePrivacy Regulation will complement the regime. Third, the current practices of the OBA sector are unlikely to be compliant with the principles of fairness, purpose limitation and data minimisation. Fourth, while some uncertainties remain, it is possible that the use of personal data in an OBA context is subject to the GDPR's regulation on profiling, which would impose even more restrictions on those uses. Throughout the four parts of the analysis, it becomes clear that the GDPR has brought the DPD's 'consent + necessity' regulatory model to a new level, bolstering both elements in different parts of the legal framework.

Such a 'consent + necessity 2.0' model was then examined critically in Chapter 5. To allow for a meaningful evaluation that takes into account the potential threats of big data flagged up in previous chapters, it was suggested that data subjects should be seen and protected as autonomous individuals,

consumers and citizens. These roles signify the various areas where personal and social preferences are of varying degrees of importance. In addition to the specific issues in the contexts of protecting autonomous, interpersonal and political interests, this regulatory model also suffers from the general rigidity of the 'libertarianism-paternalism' conceptual binary, which prevents the development of the more flexible, interest-balanced approaches.

In order to overcome these shortcomings, Chapter 6 suggested an alternative model that reconstructs the authorisation scheme of data protection law as a number of lists. By enumerating the permissive and prohibitive mechanisms that are already in place in the GDPR, this chapter shows that data protection law can be seen as comprised of a whitelist and a blacklist. Where policymakers have identified new categories of data processing that should be permitted or prohibited, they may and should consider expanding such white- or blacklists. As regards consent, there are more regulatory tools available whereby regulators may introduce different choice architectures featuring various forms of nudges and interventions. By laying down certain lists of activities that are subject to specific sets of consent requirements, policymakers are in effect creating a number of greylists of various shades. The realisation of all these regulatory possibilities by means of white-, black- and greylists will be helpful in addressing the defects of the current regime, but the adoption of such measures should be subject to the constitutional constraints in general terms.

To sum it up, a brief answer to the overarching research question would be: Considering the techno-economic realities of data uses for OBA purposes and its potential benefits and risks, data protection law should respond to the new challenges, including those posed to the individualistic, collective and societal interests, by going beyond the somewhat binary 'consent + necessity

2.0′ regulatory model as currently featured in the GDPR, and realising the available options in the regulatory toolbox.

Applicability to wider contexts

Of course, considering that the conclusion is drawn based on the investigation of a case study as well as the positive law of a specific jurisdiction, the extent to which the findings of this study can be applied to other contexts should be further clarified. To begin with, if OBA is considered one form of commercial use of big data, then the first question will naturally concern other forms of commercial use of big data. For example, it can imagined that big data is also highly valuable to the financial sector. Using sophisticated systems to analyse and predict customer behaviour for such purposes as credit scoring or insurance quoting would certainly prove efficient and accurate. Despite the slight differences in the nature of these contexts, there is no reason why the findings based on the observation of the OBA industry do not hold for the rest of the private sector. In fact, in cases where it is even clearer that personal data is involved and significant impacts are imposed on individuals, such as big data credit scoring, it would be even less disputable that the dangers are present and the legal regime is insufficient. The legal analysis would of course have to adapt to each use case of big data, but the general analytical framework should be by and large applicable to other scenarios.

Zooming out to an even greater extent, one may make further comparison with the application of big data in the public sector. While it has been highlighted in the beginning of this study that the boundaries between private and public use of data have been blurred, and that OBA as an instance of private use of big data may have public implications, it would be inadvisable to apply the findings of this study without careful adjustment to use cases that are intended for public purposes in the first place. For example, big data

surveillance has been discussed both as a possible governmental measure and as a potential threat to fundamental rights.[1] Big data for healthcare purposes, including research and development of new medicine[2] and prevention of epidemics[3], is another field where further debates will take place. These contexts, while sharing certain similarities in common, are considerably different from the case of OBA for two reasons. First, in cases of public use of big data, public interests would make a much more compelling case. It follows that, on the one hand, the legal basis of public interest, among others, would be more likely to be relied on, and on the other, policymakers would need to allocate a greater weight to this consideration when balancing against other interests. Second, commercial interests are of lesser concern in those cases, or in other words, data subjects would not be considered consumers as much as they would in the case of OBA. Therefore, the regulatory approach would face situations that have less to do with 'real choices' or 'power asymmetries', but more with 'checks and balances' or 'constitutional legitimacy'. Having said that, the general theoretical framework taking into account personal, collective and social interests will remain sound as long as the significant differences are factored in.

In an even bigger picture, it would be reasonable to discuss the transferability of the findings to emerging technologies other than big data. In fact, as highlighted in the very beginning of this study, big data has already become an integral part of many other technologies, such as the internet of things or machine learning. To the extent that the functioning of these

---

[1] For a comprehensive coverage of the literature on this topic, see David Lyon, 'Big Data Surveillance: Snowden Everyday Practices and Digital Futures' in Tugba Basaran and others (eds), *International Political Sociology: Transversal Lines* (Routledge 2017).
[2] Doug Howe and others, 'The Future of Biocuration' (2008) 455(4) Nature 47.
[3] Stephen J. Mooney, Daniel J. Westreich and Abdulrahman M. El-Sayed, 'Epidemiology in the Era of Big Data' (2015) 26(3) Epidemiology.

technologies is dependent on powerful analysis of large amounts of personal data, the doctrinal and theoretical frameworks employed in this study would likely be able to accommodate such technologies as well.

In terms of jurisdiction, while the legal analysis throughout the previous chapters has been overwhelmingly focused on the EU's GDPR, most of the non-doctrinal observations would nevertheless go for other jurisdictions both intra- and extra-EU. For Member States of the EU, while the GDPR has achieved a greater level of harmonisation, a number of issues remain at the discretion of national laws.[4] For example, and as has been highlighted in the previous chapter, national legislation may blacklist certain uses of sensitive data. This would leave a role, though limited, for the policymakers or regulators of Member States to continue to play. The policymaking of Member States would be subject to the general analytical framework here, as well as the constitutional constraints suggested. For non-EU jurisdictions, the 'black-letter' part of this study would of course not apply, but apart from that, the big data risks they are facing will be largely similar, and thus the general approach, at least regarding the various possibilities in the regulatory toolbox, should be equally applicable to these jurisdictions. The specific implementation may vary from one country to another, depending on their legal system and legal culture, but the overarching ideas advanced by this study should be able to more or less inform policymakers across jurisdictions.

Limitations and future research

The conclusion of this study should not be seen as only the end of an inquiry, but rather the beginning of longer-term academic and public debates. There

[4] For a full list of these derogations, see Jiahong Chen, 'How the Best-Laid Plans Go Awry: The (Unsolved) Issues of Applicable Law in the General Data Protection Regulation' (2016) 6(4) International Data Privacy Law 310.

are many interesting areas regarding big data and data protection law that cannot be covered by this study, but should be addressed in future research. It is impossible to name all the potentially relevant research topics exhaustively here, but highlighting some of the unexplored fields should be helpful in sparkling ideas about further research initiatives.

Some of the interesting topics are not covered mainly due to the choice of methodologies within this study. Empirical evidence is needed to further uncover the industrial practices, as well as the inferred ramifications, such as the hypothetical network effects of the dominant players in the ecosystem, as suggested in Chapter 1. Impartial, well-designed surveys on the overall attitudes to and preferences of certain categories of use of personal data, as informed by the discussion in Chapter 2, are also needed, in particular after the recent disclosures of how Internet users' online data might have been misused.

The theories about the risks arising from big data also need further support, both empirically and theoretically. Certain ideas about the non-commercial collective interests, for instance, have been lightly touched upon in Chapter 3, but will need to be fleshed out with more in-depth discussions. Also, while the choice of a 'risk-based' framework to conceptualise the impact of big data is useful for the purpose of this study, it is also possible to explore how big data affects certain public goods in ways that do not necessarily involve any short-term or long-term risks. If we consider the control over the use of personal data as a collective decision on how we govern a society at present and in the future, the idea of 'good (or better) governance' itself would not have to have recourse to the establishment of any harm.

As regards the legal part of the research, it has been well acknowledged that there are many aspects of the GDPR that have not been addressed by

Chapter 4. Some new provisions (especially the procedural ones) that might actually make a difference, such as data protection by design[5] and data protection impact assessments[6], are yet to be fully discussed in the context of OBA. The limitations of the current legal framework, as the subject matter of Chapter 5, have been highlighted, but are not considered in detail with regard to the potential interactions with other areas of law, including consumer protection law, competition law, media law or even election law.

Much work will also be needed to further theorise the list-based approach as proposed in Chapter 6, where a general framework is provided but it has to be filled up with more substance. The activities that should be put on the white-/grey-/blacklists would of course be the most important issues to consider if that is the approach taken by the policymakers. Equally important are the procedural and substantive constraints that should be imposed on policymakers and regulators in the course of codifying and enforcing such lists. In more general terms, since both big data and data protection law are developing and interacting with each other, further research on policy learning in this context is also needed.

The list of potential research topics can go on and on. The scope of the relevant yet unsolved issues shows how much it would take to fully address the intractable problems surrounding data protection law and big data — which can be an ambitious, lifetime project. We have been, and will probably continue to be, in a time where big data exerts great influence on personal, communal and public life. Perhaps the good news is that regulatory and public awareness of the latent impact of big data seems to be on the rise. While the

[5] 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] OJ L119/1 ('GDPR'), art 25.
[6] ibid art 35.

aim of this study is not to provide a comprehensive design of a model data protection statute, it is hoped that the discussions here offer a helpful conceptual framework for policymakers to consider, and for the whole society to debate, on how our data protection law should respond to the looming challenges of big data.

Bibliography

Primary Sources

Case-law

BVerfGE 65, 1 [1965]
Durant v FSA [2003] EWCA Civ 1746
Joined Cases C-92/09 and C-93/09 *Volker und Markus Schecke and Eifert* [2010] OJ C 13/6
Joined cases C-468/10 and C-469/10 *ASNEF* [2011] OJ C 25/18
Case C-70/10 *Scarlet Extended* [2012] OJ C 25/6
Joined Cases C-293/12 and C-594/12 *Digital Rights Ireland and Seitlinger and Others* [2014] OJ C 175/6
Case C-131/12 *Google Spain and Google* [2014] OJ C 212/4
Case C-362/14 *Schrems* [2015] OJ C 398/5
Case C-582/14 *Breyer* [2016] OJ C 475/3
DB v GMC [2016] EWHC 2331 (QB)
Joined Cases C-203/15 and C-698/15 *Tele2 Sverige* [2016] OJ C 53/11

Legislation

Data Protection Act 1998 ('Data Protection Act 1998')
European Convention on Human Rights ('ECHR')
Hessisches Datenschutzgesetz [1970]
Council of Europe Committee of Ministers Resolution (73)22 on the Protection of the Privacy of Individuals vis-à-vis Electronic Data Banks in the Private Sector ('Resolution 73(22)') [1973]
Datalag (of Sweden) [1973]
Decreto de Aprovação da Constituição da República Portuguesa [1976]
Bundesdatenschutzgesetz (of Germany) [1977]
Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés (of France) [1978]
OECD Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data ('OECD Guidelines') [1980]
Council of Europe Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data [1981] ETS No. 108
Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts ('UCTD') [1993] OJ L95/29
Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data ('DPD') [1995] OJ L281/31
Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) (As amended by Directive 2009/136/EC) ('ePrivacy Directive') [2002] OJ L201/37
Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and

amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council ('Unfair Commercial Practices Directive') ('UCPD') [2005] OJ L149/22

Treaty of Lisbon amending the Treaty on European Union and the Treaty establishing the European Community ('Treaty of Lisbon') [2007] OJ C306/1

Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws [2009] OJ L337/11

Charter of Fundamental Rights of the European Union ('Charter') [2012] OJ C326/391

Consolidated version of the Treaty on European Union ('TEU') [2012] OJ C326/13

Consolidated version of the Treaty on the Functioning of the European Union ('TFEU') [2012] OJ C326/47

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) ('GDPR') [2016] OJ L119/1

Secondary Sources

Books

Bainbridge D, *Data Protection* (2nd edn, XPL Law 2005)

Battelle J, *The Search: How Google and Its Rivals Rewrote the Rules of Business and Transformed Our Culture* (Nicholas Brealey Publishing 2005)

Bennett CJ and Raab CD, *The Governance of Privacy: Policy Instruments in Global Perspective* (The MIT Press 2006)

Bernal P, *Internet Privacy Rights: Rights to Protect Autonomy* (Cambridge University Press 2014)

Brown I and Marsden CT, *Regulating Code : Good Governance and Better Regulation in the Information Age* (The MIT Press 2013)

Bygrave LA, *Data Privacy Law: An International Perspective* (Oxford University Press 2014)

Cleland S, *Search & Destroy: Why You Can't Trust Google Inc.* (Telescope Books 2011)

Craig T and Ludloff ME, *Privacy and Big Data* (O'Reilly 2011)

Dempster C and Lee J, *The Rise of the Platform Marketer: Performance Marketing with Google, Facebook, and Twitter, Plus the Latest High-Growth Digital Advertising Platforms* (John Wiley & Sons, Inc. 2015)

Dworkin G, *The Theory and Practice of Autonomy* (Cambridge University Press 1988)

Dworkin R, *Sovereign Virtue: The Theory and Practice of Equality* (Harvard University Press 2000)

González Fuster G, *The Emergence of Personal Data Protection as a Fundamental Right of the EU* (Springer 2014)

Greenleaf G, *Asian Data Privacy Laws: Trade and Human Rights Perspectives* (Oxford University Press 2014)

Hartzog W, *Privacy's Blueprint: The Battle to Control the Design of New Technologies* (Harvard University Press 2018)

Hijmans H, *The European Union as Guardian of Internet Privacy: The Story of Art 16 TFEU* (Springer 2016)
Hildebrandt M, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology* (Edward Elgar 2015)
John P, *How Far to Nudge? Assessing Behavioural Public Policy* (Edward Elgar 2018)
Kant I, *Groundwork for the Metaphysics of Morals* (Abbott TK tr, Broadview 2005)
Kirkpatrick D, *The Facebook Effect: The Inside Story of the Company That Is Connecting the World* (Simon & Schuster 2010)
Korff D, *Data Protection Law in Practice in the European Union* (Federation of European Direct and Interactive Marketing and The Direct Marketing Association 2005)
Kosta E, *Consent in European Data Protection Law* (Brill 2013)
Kreiss D, *Prototype Politics: Technology-intensive Campaigning and the Data of Democracy* (Oxford University Press 2016)
Krotoszynski RJ, Jr., *Privacy Revisited: A Global Perspective on the Right to Be Left Alone* (Oxford University Press 2016)
Lessig L, *Code: Version 2.0* (2nd edn, Basic Books 2006)
Levy S, *In the Plex: How Google Thinks, Works, and Shapes Our Lives* (Simon & Schuster 2011)
Lynskey O, *The Foundations of EU Data Protection Law* (Oxford University Press 2015)
Lyon D, *The Electronic Eye: The Rise of Surveillance Society* (University of Minnesota Press 1994)
Makulilo AB (ed), *African Data Privacy Laws* (Springer 2016)
Marichal J, *Facebook Democracy: The Architecture of Disclosure and the Threat to Public Life* (Routledge 2016)
Mayer-Schönberger V, *Delete: The Virtue of Forgetting in the Digital Age* (Princeton University Press 2009)
Mayer-Schönberger V and Cukier K, *Big Data: A Revolution That Will Transform How We Live, Work and Think* (John Murray 2013)
Nissenbaum H, *Privacy in Context: Technology, Policy, and the Integrity of Social Life* (Stanford University Press 2010)
Ohlhorst F, *Big Data Analytics: Turning Big Data into Big Money* (John Wiley & Sons, Inc. 2013)
Pariser E, *The Filter Bubble: What the Internet Is Hiding from You* (The Penguin Press 2011)
Pasquale F, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Harvard University Press 2015)
Rawls J, *A Theory of Justice* (revised edn, Harvard University Press 2009)
Raz J, *The Morality of Freedom* (Oxford University Press 1986)
——, *Between Authority and Interpretation: On the Theory of Law and Practical Reason* (Oxford University Press 2009)
Regan PM, *Legislating Privacy: Technology, Social Values, and Public Policy* (University of North Carolina Press 1995)
Ruiz BR, *Privacy in Telecommunications: A European and an American Approach* (Kluwer Law International 1997)
Siegel E, *Predictive Analytics : The Power to Predict Who Will Click, Buy, Lie, or Die* (Wiley 2016)
Smith M, *Targeted: How Technology Is Revolutionizing Advertising and the Way Companies Reach Consumers* (American Management Association 2015)
Solove DJ, *Understanding Privacy* (Harvard University Press 2008)
——, *Nothing to Hide: The False Tradeoff Between Privacy and Security* (Yale University Press 2011)
Stross R, *Planet Google: One Company's Audacious Plan to Organize Everything We Know* (Free Press 2008)

Sumner S, *You: For Sale: Protecting Your Personal Data and Privacy Online* (Elsevier 2016)
Sumpter D, *Outnumbered: From Facebook and Google to Fake News and Filter-bubbles – The Algorithms That Control Our Lives* (Bloomsbury 2018)
Sunstein CR, *Republic.com 2.0* (Princeton University Press 2007)
——, *Why Nudge? The Politics of Libertarian Paternalism* (Yale University Press 2014)
——, *The Ethics of Influence: Government in the Age of Behavioral Science* (Cambridge University Press 2016)
Taebi B and Roeser S (eds), *The Ethics of Nuclear Energy: Risk, Justice, and Democracy in the post-Fukushima Era* (Cambridge University Press 2015)
Thaler RH and Sunstein CR, *Nudge: Improving Decisions About Health, Wealth, and Happiness* (Yale University Press 2008)
Turow J, *The Daily You: How the New Advertising Industry is Defining Your Identity and Your World* (Yale University Press 2012)
Vaidhyanathan S, *The Googlization of Everything (and Why We Should Worry)* (University of California Press 2011)
Van Dijk J, *The Network Society* (3rd edn, SAGE 2012)
Weatherill S, *EU Consumer Law and Policy* (Elgar 2013)
Westin AF, *Privacy and Freedom* (The Bodley Head 1967)
Wresch W, *Disconnected: Haves and Have-Nots in the Information Age* (Rutgers University Press 1996)
Zuiderveen Borgesius FJ, *Improving Privacy Protection in the Area of Behavioural Targeting* (Kluwer Law International 2015)

Contributions to Edited Books

Carolan E and Spina A, 'Behavioural Sciences and EU Data Protection Law: Challenges and Opportunities' in Alemanno A and Sibony A-L (eds), *Nudge and the Law: A European Perspective* (Hart Publishing 2015)
Cassese S, 'Exploring the Legitimacy of Nudging' in Kemmerer A and others (eds), *Choice Architecture in Democracies : Exploring the Legitimacy of Nudging* (Nomos 2016)
Castelluccia C, Kaafar M-A and Tran M-D, 'Betrayed by Your Ads! Reconstructing User Profiles from Targeted Ads' in Fischer-Hübner S and Wright M (eds), *Privacy Enhancing Technologies* (Springer 2012)
Chester J, 'Cookie Wars: How New Data Profiling and Targeting Techniques Threaten Citizens and Consumers in the "Big Data" Era' in Gutwirth S and others (eds), *European Data Protection: In Good Health?* (Springer 2012)
Dowling G and Weeks W, 'Measuring Media Corporate Reputations' in Helm S, Liehr-Gobbers K and Storck C (eds), *Reputation Management* (2011)
Epstein RA, 'Contracts Small and Contract Large: Contract Law through the Lens of Laissez-Faire' in Buckley FH (ed), *The Fall and Rise of Freedom of Contract* (Duke University Press 1999)
Greenleaf G, 'Data Protection in a Globalised Network' in Brown I (ed), *Research Handbook on Governance of the Internet* (Edward Elgar 2013)
Heinig HM, 'Autonomy vs. Technocracy: Libertarian Paternalism Revisited' in Kemmerer A and others (eds), *Choice Architecture in Democracies : Exploring the Legitimacy of Nudging* (Nomos 2016)
Hempel L and Lammerant H, 'Impact Assessments as Negotiated Knowledge' in Gutwirth S, Leenes R and De Hert P (eds), *Reforming European Data Protection Law* (Springer 2015)

Kammourieh L and others, 'Group Privacy in the Age of Big Data' in Taylor L, Floridi L and Van der Sloot B (eds), *Group Privacy: New Challenges of Data Technologies* (Springer 2017)
Lübbe-Wolff G, 'Constitutional Limits to Health-Related Nudging – a Matter of Balancing' in Kemmerer A and others (eds), *Choice Architecture in Democracies : Exploring the Legitimacy of Nudging* (Nomos 2016)
Lyon D, 'Big Data Surveillance: Snowden Everyday Practices and Digital Futures' in Basaran T and others (eds), *International Political Sociology: Transversal Lines* (Routledge 2017)
Mantelero A, 'From Group Privacy to Collective Privacy: Towards a New Dimension of Privacy and Data Protection in the Big Data Era' in Taylor L, Floridi L and Van der Sloot B (eds), *Group Privacy: New Challenges of Data Technologies* (Springer 2017)
McCrudden C and King J, 'The Dark Side of Nudging: The Ethics, Political Economy, and Law of Libertarian Paternalism' in Kemmerer A and others (eds), *Choice Architecture in Democracies : Exploring the Legitimacy of Nudging* (Nomos 2016)
Mendoza I and Bygrave LA, 'The Right Not to be Subject to Automated Decisions Based on Profiling' in Synodinou T-E and others (eds), *EU Internet Law: Regulation and Enforcement* (2017)
Quelle C, 'Not just User Control in the General Data Protection Regulation: On the Problems with Choice and Paternalism, and on the Point of Data Protection' in Lehmann A and others (eds), *Privacy and Identity Management: Facing up to Next Steps* (Springer 2016)
Rouvroy A and Poullet Y, 'The Right to Informational Self-Determination and the Value of Self-Development: Reassessing the Importance of Privacy for Democracy' in Gutwirth S and others (eds), *Reinventing Data Protection?* (Springer 2009)
Stahl BC, 'The Impact of the UK Human Rights Act 1998 on Privacy Protection in the Workplace' in Subramanian R (ed), *Computer Security, Privacy and Politics: Current Issues, Challenges, and Solutions* (IRM Press 2008)
Stilgoe J, 'Foreword: Why Responsible Innovation?' in Owen R, Bessant J and Heintz M (eds), *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society* (Wiley 2013)
Sugarman J and Sipp D, 'Translational Stem Cell Research: Issues Beyond the Debate on the Moral Status of the Human Embryo' in Hug K and Hermerén G (eds), *Stem Cell Biology and Regenerative Medicine* (Humana Press 2011)
Sunstein CR, 'The Ethics of Choice Architecture' in Kemmerer A and others (eds), *Choice Architecture in Democracies : Exploring the Legitimacy of Nudging* (Nomos 2016)
Tavani HT, 'Privacy and Security' in Langford D (ed), *Internet Ethics* (Palgrave Macmillan 2000)
van Aaken A, 'Constitutional Limits to Paternalistic Nudging: A Proportionality Assessment' in Kemmerer A and others (eds), *Choice Architecture in Democracies : Exploring the Legitimacy of Nudging* (Nomos 2016)
Volkmann U, 'Nudging, Education, Paternalism: A Philosophical Perspective from the Old Europe' in Kemmerer A and others (eds), *Choice Architecture in Democracies : Exploring the Legitimacy of Nudging* (Nomos 2016)
Zuiderveen Borgesius F, 'Behavioural Sciences and the Regulation of Privacy on the Internet' in Alemanno A and Sibony A-L (eds), *Nudge and the Law: A European Perspective* (Hart Publishing 2015)

Journal Articles

Abadie A and Gay S, 'The Impact of Presumed Consent Legislation on Cadaveric Organ Donation: A Cross-country Study' (2006) 25(4) Journal of Health Economics

Akerlof GA, 'The Market for "Lemons": Quality Uncertainty and the Market Mechanism' (1970) 84(3) The Quarterly Journal of Economics 488
Awad NF and Krishnan MS, 'The Personalization Privacy Paradox: An Empirical Evaluation of Information Transparency and the Willingness to Be Profiled Online for Personalization' (2006) 30(1) MIS Quarterly 13
Baldwin-Philippi J, 'The Myths of Data-Driven Campaigning' (2017) 34 Political Communication 627
Bedi M, 'The Curious Case of Cell Phone Location Data: Fourth Amendment Doctrine Mash-up' (2016) 110(2) Northwestern University Law Review 507
Ben-Shahar O and Schneider CE, 'The Failure of Mandated Disclosure' (2011) 159 University of Pennsylvania Law Review 647
Benn SI, 'Freedom, Autonomy and the Concept of a Person' (1975) 76 Proceedings of the Aristotelian Society 109
Bennett CJ, 'Voter Databases, Micro-targeting, and Data Protection Law: Can Political Parties Campaign in Europe as They Do in North America?' (2016) 6(4) International Data Privacy Law
Bennett SC, 'Regulating Online Behavioral Advertising' (2011) 44 The John Marshall Law Review 899
Bergemann D and Bonatti A, 'Selling Cookies' (2015) 7(3) American Economic Journal: Microeconomics 259
Bodó B, Helberger N and De Vreese CH, 'Political Micro-targeting: A Manchurian Candidate or Just a Dark Horse?' (2017) 6(4) Internet Policy Review 3
Boerman SC, Kruikemeier S and Zuiderveen Borgesius FJ, 'Online Behavioral Advertising: A Literature Review and Research Agenda' (2017) 46(3) Journal of Advertising 363
Boyd D and Crawford K, 'Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon' (2012) 15(5) Information, Communication & Society 662
Bozdag E, 'Bias in Algorithmic Filtering and Personalization' (2013) 15 Ethics and Information Technology 209
Chen J, 'How the Best-Laid Plans Go Awry: The (Unsolved) Issues of Applicable Law in the General Data Protection Regulation' (2016) 6(4) International Data Privacy Law 310
——, 'The Dangers of Accuracy: Exploring the Other Side of the Data Quality Principle' (2018) 4(1) European Data Protection Law Review 36
Citron DK and Pasquale F, 'The Scored Society: Due Process for Automated Predictions' (2014) 89 Washington Law Review 1
Cofone IN, 'The Way the Cookie Crumbles: Online Tracking Meets Behavioural Economics' [2017] International Journal of Law and Information Technology
Cohen JE, 'Examined Lives: Informational Privacy and the Subject as Object' (2000) 52 Stanford Law Review 1373
——, 'What Privacy Is For' (2013) 126 Harvard Law Review 1904
Cranor LF, 'Necessary But Not Sufficient: Standardized Mechanisms for Privacy Notice and Choice' (2012) 10 Journal on Telecommunications and High Technology Law 273
Crawford K and Schultz J, 'Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms' (2014) 55 Boston College Law Review 93
Danaher J, 'The Threat of Algocracy: Reality, Resistance and Accommodation' (2016) 29(3) Philosophy & Technology 245
Edwards L, 'Taking the "Personal" Out of Personal Data: *Durant v FSA* and its Impact on the Legal Regulation of CCTV' (2004) 1(2) SCRIPT-ed 341

Edwards L and Veale M, 'Slave to the Algorithm? Why a "Right to an Explanation" Is Probably Not the Remedy You Are Looking For' (2017) 16(1) Duke Law and Technology Review 1
Eijk Nv and others, 'Online Tracking: Questioning the Power of Informed Consent' (2012) 14(5) Info 57
Fairfield JAT, 'Do-Not-Track as Default' (2013) 11(7) Northwestern Journal of Technology and Intellectual Property 575
Fairfield JAT and Engel C, 'Privacy as a Public Good' (2015) 65 Duke Law Journal
Feeny D and others, 'The Tragedy of the Commons: Twenty-two Years Later' (1990) 18(1) Human Ecology 1
Fotion N, 'Paternalism' (1979) 89(2) Ethics 191
Fried C, 'Privacy' (1968) 77(3) The Yale Law Journal 475
Goldfarb A and Tucker CE, 'Privacy Regulation and Online Advertising' (2011) 57(1) Management Science 57
Grimmer J, 'We Are All Social Scientists Now: How Big Data, Machine Learning, and Causal Inference Work Together' (2015) 48(1) PS: Political Science & Politics 80
Hacker P, 'Personal Data, Exploitative Contracts, and Algorithmic Fairness: Autonomous Vehicles Meet the Internet of Things ' (2017) 7(4) International Data Privacy Law 1
Hausman DM and Welch B, 'Debate: To Nudge or Not to Nudge' (2010) 18(1) The Journal of Political Philosophy
Hermstrüwer Y, 'Contracting Around Privacy: The (Behavioral) Law and Economics of Consent and Big Data' (2017) 7(3) JIPITEC 9
Hermstrüwera Y and Dickert S, 'Sharing is Daring: An Experiment on Consent, Chilling Effects and a Salient Privacy Nudge' (2017) 51 International Review of Law and Economics 38
Holland HB, 'Privacy Paradox 2.0' (2010) 19 Widener Law Journal 893
Hoofnagle CJ and Urban JM, 'Alan Westin's Privacy *Homo Economicus*' (2014) 49 Wake Forest Law Review 261
Hoofnagle CJ and Whittington J, 'Free: Accounting for the Costs of the Internet's Most Popular Price' (2014) 61 UCLA Law Review 606
Howe D and others, 'The Future of Biocuration' (2008) 455(4) Nature 47
Issenberg S, 'How President Obama's Campaign Used Big Data to Rally Individual Voters' (2013) 118(1) MIT Technology Review 38
Jensen C, Potts C and Jensen C, 'Privacy Practices of Internet Users: Self-reports versus Observed Behavior' (2005) 63 International Journal of Human-Computer Studies 203
Johnson EJ and Goldstein D, 'Do Defaults Save Lives?' (2003) 302 Science 1338
Kahneman D, 'Maps of Bounded Rationality: Psychology for Behavioral Economics' (2003) 93(5) The American Economic Review 1449
Kamara I and Kosta E, 'Do Not Track Initiatives: Regaining the Lost User Control' (2016) 6(4) International Data Privacy Law 276
Kerber W, 'Digital Markets, Data, and Privacy: Competition Law, Consumer Law and Data Protection ' (2016) 11 Journal of Intellectual Property Law & Practice 856
Kirkup G, 'Academic Blogging: Academic Practice and Academic Identity' (2010) 8(1) London Review of Education 75
Kokott J and Sobotta C, 'The Distinction Between Privacy and Data Protection in the Jurisprudence of the CJEU and the ECtHR' (2013) 3(4) International Data Privacy Law 222
Koops B-J, 'The Trouble with European Data Protection Law' (2014) 4(4) International Data Privacy Law 250

Koops B-J and Leenes R, 'Privacy Regulation Cannot Be Harcoded. A Critical Comment on the "Privacy by Design" Provision in Data-Protection Law' (2014) 28(2) International Review of Law, Computers & Technology 159
Kosta E, 'Peeking into the Cookie Jar: The European Approach Towards the Regulation of Cookies' (2013) 21(4) International Journal of Law and Information Technology 380
Kostić B and Vargas Penagos E, 'The Freely Given Consent and the "Bundling" Provision under the GDPR' (2017) 153 Computerrecht 217
Lazaro C and Le Métayer D, 'Control over Personal Data: True Remedy or Fairy Tale?' (2015) 12(1) SCRIPT-ed 3
Lazer D and others, 'The Parable of Google Flu: Traps in Big Data Analysis' (2014) 343 Science 1203
Leenes R, 'Do They Know Me? Deconstructing Identifiability' (2007) 4(1&2) University of Ottawa Law & Technology Journal 135
Lenard TM and Rubin PH, 'In Defense of Data: Information and the Costs of Privacy' (2010) 2(1) Policy & Internet 149
Lessig L, 'The Law of the Horse: What Cyberlaw Might Teach' (1999) 113 Harvard Law Review 501
Li F, Williams H and Bogle M, 'The "Millennium Bug": Its Origin, Potential Impact and Possible Solutions' (1999) 19(1) International Journal of Information Management 3
Lorber S, 'Data Protection: A contextual Approach to Regulation' (2014) 14(5) Privacy and Data Protection 11
Luzak JA, 'Privacy Notice for Dummies? Towards European Guidelines on How to Give 'Clear and Comprehensive Information' on the Cookies' Use in Order to Protect the Internet Users' Right to Online Privacy' (2014) 37 Journal of Consumer Policy 547
Lynskey O, 'Deconstructing Data Protection: The "Added-value" of a Right to Data Protection in the EU Legal Order' (2014) 63 International and Comparative Law Quarterly 569
——, 'Aligning Data Protection Rights with Competition Law Remedies? The GDPR Right to Data Portability' (forthcoming) European Law Journal
MacCarthy M, 'New Directions in Privacy: Disclosure, Unfairness and Externalities' (2011) 6(3) I/S: A Journal of Law and Policy 425
Mantelero A, 'Competitive Value of Data Protection: The Impact of Data Protection Regulation on Online Behaviour' (2013) 3(4) International Data Privacy Law 229
——, 'The Future of Consumer Data Protection in the E.U.: Re-thinking the "Notice and Consent" Paradigm in the New Era of Predictive Analytics' (2014) 30 Computer Law & Security Review 643
Mayer-Schönberger V and Padova Y, 'Regime Change? Enabling Big Data Through Europe's New Data Protection Regulation' (2016) XVII The Columbia Science & Technology Law Review 315
McCullagh K, 'Protecting "Privacy" Through Control of "Personal" Data Processing: A Flawed Approach' (2009) 23(1-2) International Review of Law, Computers & Technology 13
McDonald AM and Cranor LF, 'The Cost of Reading Privacy Policies' (2008) 4(3) I/S: A Journal of Law and Policy 543
McLaughlin MM and Vaupel S, 'Constitutional Right of Privacy and Investigative Consumer Reports: Little Brother Is Watching You' (1975) 2 Hastings Constitutional Law Quarterly 773

McNamara K, 'The Paparazzi Industry and New Media: The Evolving Production and Consumption of Celebrity News and Gossip Websites' (2011) 14(5) International Journal of Cultural Studies 515
McStay A, 'I Consent: An Analysis of the Cookie Directive and Its Implications for UK Behavioral Advertising' (2012) 15(4) New Media & Society 596
——, 'Empathic Media and Advertising: Industry, Policy, Legal and Citizen Perspectives (the Case for Intimacy)' [2016](July-December) Big Data & Society 1
Merrill TW and Smith HE, 'Optimal Standardization in the Law of Property: The *Numerus Clausus* Principle' (2000) 110(1) The Yale Law Journal 1
Millard C and Hon WK, 'Defining "Personal Data" in E-social Science' (2012) 15(1) Information, Communication & Society 66
Miller AA, 'What Do We Worry About When We Worry About Price Discrimination? The Law and Ethics of Using Personal Information for Pricing' (2014) 19 Journal of Technology Law and Policy 41
Mooney SJ, Westreich DJ and El-Sayed AM, 'Epidemiology in the Era of Big Data' (2015) 26(3) Epidemiology
Moor JH, 'Towards a Theory of Privacy in the Information Age' (1997) 27(3) ACM SIGCAS Computers and Society 27
Moreham NA, 'Privacy in Public Places' (2006) 65(3) Cambridge Law Journal 606
Mostert M and others, 'From Privacy to Data Protection in the EU: Implications for Big Data Health Research' (2017) 25(1) European Journal of Health Law 43
Newman N, 'Search, Antitrust, and the Economics of the Control of User Data' (2014) 31(2) Yale Journal on Regulation 401
Norberg PA, Horne DR and Horne DA, 'The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors' (2007) 41(1) The Journal of Consumer Affairs 100
Noto La Diega G, 'Some Considerations on Intelligent Online Behavioural Advertising' (2018) 66-67 Revue du droit des technologies de l'information 53
Nozick R, 'Invisible-Hand Explanations' (1994) 84(2) The American Economic Review 314
O'Keeffe GS, Clarke-Pearson K and Council on Communications and Media, 'Clinical Report—The Impact of Social Media on Children, Adolescents, and Families' (2011) 127(4) Pediatrics 800
O'Leary DE, 'Artificial Intelligence and Big Data' (2013) 28(2) IEEE Intelligent Systems 96
——, '"Big Data", the "Internet of Things" and the "Internet of Signs"' (2013) 20(1) Intelligent Systems in Accounting, Finance and Management 53
Ohm P, 'The Rise and Fall of Invasive ISP Surveillance' [2009] University of Illinois Law Review 1417
——, 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' (2010) 57 UCLA Law Review 1701
Oostveen M, 'Identifiability and the Applicability of Data Protection to Big Data' (2016) 6(4) International Data Privacy Law 299
Posner RA, 'The Right of Privacy' (1978) 12 Georgia Law Review 393
Prosser WL, 'Privacy' (1960) 48(3) California Law Review 383
Rauhofer J, 'Privacy Is Dead, Get Over It! Information Privacy and the Dream of a Risk-free Society' (2008) 17(3) Information & Communications Technology Law 185
——, 'Of Men and Mice: Should the EU Data Protection Authorities' Reaction to Google's New Privacy Policy Raise Concern for the Future of the Purpose Limitation Principle?' (2015) 1(1) European Data Protection Law Review 5
Rothschild M and Stiglitz J, 'Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information' (1976) 90(4) The Quarterly Journal of Economics 629

Rubinstein IS, 'Big Data: The End of Privacy or a New Beginning?' (2013) 3(2) International Data Privacy Law 74
Sableman M, Shoenberger H and Thorson E, 'Consumer Attitudes Toward Relevant Online Behavioral Advertising: Crucial Evidence in the Data Privacy Debates' (2013) 2013(1) Media Law Resource Center Bulletin 93
Schermer BW, Custers B and Hof Svd, 'The Crisis of Consent: How Stronger Legal Protection May Lead to Weaker Consent in Data Protection' (2014) 16 Ethics and Information Technology 171
Schwartz PM, 'Privacy and Democracy in Cyberspace' (1999) 52 Vanderbilt Law Review 1609
——, 'Internet Privacy and the State' (2000) 32 Connecticut Law Review 815
Schwartz PM and Solove DJ, 'The PII Problem: Privacy and a New Concept of Personally Identifiable Information' (2011) 86 New York University Law Review 1814
Sharpe SA, 'Asymmetric Information, Bank Lending, and Implicit Contracts: A Stylized Model of Customer Relationships' (1990) XLV(4) The Journal of Finance 1069
Simmonds NE, 'Property, Autonomy and Welfare' (1981) 67(1) ARSP 61
Smit EG, Noort GV and Voorveld HAM, 'Understanding Online Behavioural Advertising: User Knowledge, Privacy Concerns and Online Coping Behaviour in Europe' (2014) 32 Computers in Human Behavior 15
Solove DJ, 'Privacy and Power: Computer Databases and Metaphors for Information Privacy' (2001) 53 Stanford Law Review 1393
——, 'Privacy Self-Management and the Consent Dilemma' (2013) 126 Harvard Law Review 1880
Sovern J, 'Opting in, Opting out, or No Options at All: The Fight for Control of Personal Information' (1999) 74 Washington Law Review 1033
Spence M, 'Job Market Signaling' (1973) 87(3) The Quarterly Journal of Economics 355
Spina A, 'Risk Regulation of Big Data: Has the Time Arrived for a Paradigm Shift in EU Data Protection Law?' (2014) 2 European Journal of Risk Regulation 248
Stilgoe J, Owen R and Macnaghten P, 'Developing a Framework for Responsible Innovation' (2013) 42 Research Policy 1568
Summers CA, Smith RW and Reczek RW, 'An Audience of One: Behaviorally Targeted Ads as Implied Social Labels' (2016) 43 Journal of Consumer Research 156
Sunstein CR, 'The Storrs Lectures: Behavioral Economics and Paternalism' (2013) 122 The Yale Law Journal 1826
Taebi B and others, 'Responsible Innovation as an Endorsement of Public Values: The Need for Interdisciplinary Research ' (2014) 1 Journal of Responsible Innovation 118
Tene O, 'What Google Knows: Privacy and Internet Search Engines' [2008](4) Utah Law Review 1433
Tene O and Polonetsky J, 'To Track or "Do Not Track": Advancing Transparency and Individual Control in Online Behavioural Advertising' (2012) 13(1) Minnesota Journal of Law, Science & Technology 281
——, 'Big Data for All: Privacy and User Control in the Age of Analytics' (2013) 11(5) Northwestern Journal of Technology and Intellectual Property 239
van der Sloot B, 'Decisional Privacy 2.0: The Procedural Requirements Implicit in Article 8 ECHR and Its Potential Impact on Profiling' (2017) 7(3) International Data Privacy Law 190
Warren SD and Brandeis LD, 'The Right to Privacy' (1890) IV(5) Harvard Law Review 193
West R, 'Authority, Autonomy, and Choice: The Role of Consent in the Moral and Political Visions of Franz Kafka and Richard Posner' (1985) 99(2) Harvard Law Review 384

Zarsky TZ, 'Understanding Discrimination in the Scored Society' (2014) 89 Washington Law Review 1375
——, 'Incompatible: The GDPR in the Age of Big Data' (2016) 47 Seton Hall Law Review 995
Zingales N, 'Between a Tock and Two Hard Places: WhatsApp at the Crossroad of Competition, Data protection and Consumer law' (2017) 33(4) Computer Law & Security Review 553
Zuiderveen Borgesius F and Poort J, 'Online Price Discrimination and EU Data Privacy Law' (2017) 40(3) Journal of Consumer Policy
Zuiderveen Borgesius FJ, 'Personal Data Processing for Behavioural Targeting: Which Legal Basis?' (2015) 5(3) International Data Privacy Law 163
Zuiderveen Borgesius FJ and others, 'Tracking Walls, Take-It-Or-Leave-It Choices, the GDPR, and the ePrivacy Regulation' (2017) 3(3) European Data Protection Law Review 353

Online Articles

Clifford D, 'Citizen-consumers in a Personalised Galaxy: Emotion Influenced Decision-making, a True Path to the Dark Side?' 31/2017 CiTiP Working Paper <http://ssrn.com/abstract=3037425> accessed 21 February 2018
Clifford D and Ausloos J, 'Data Protection and the Role of Fairness' 29/2017 CiTiP Working Paper Series <https://ssrn.com/abstract=3013139> accessed 14 December 2017
Cohen JE, 'Turning Privacy Inside Out' <https://ssrn.com/abstract=3162178> accessed 16 May 2018
Hammock M and Rubin PH, 'Applications Want to Be Free: Privacy Against Information' <http://ssrn.com/abstract=1781906> accessed 3 May 2018
Jernigan C and Mistree BFT, 'Gaydar: Facebook Friendships Expose Sexual Orientation' 14 First Monday <http://firstmonday.org/ojs/index.php/fm/article/viewArticle/2611> accessed 7 January 2017
Nithyanand R and others, 'Adblocking and Counter-Blocking: A Slice of the Arms Race' <https://arxiv.org/abs/1605.05077v2> accessed 12 December 2016
Rauhofer J, 'One Step Forward, Two Steps Back? Critical Observations on the Proposed Reform of the EU Data Protection Framework' 2013/17 University of Edinburgh School of Law Research Paper Series <http://ssrn.com/abstract=2260967> accessed 24 October 2014
——, 'Round and Round the Garden? Big Data, Small Government and the Balance of Power in the Information Age' 2014/06 University of Edinburgh School of Law Research Paper Series <http://ssrn.com/abstract=2389981> accessed 22 October 2014
Soltani A and others, 'Flash Cookies and Privacy' <https://ssrn.com/abstract=1446862> accessed 7 January 2017
Turow J and others, 'Contrary to What Marketers Say, Americans Reject Tailored Advertising and Three Activities That Enable It' <http://ssrn.com/abstract=1478214> accessed 20 February 2015

Conference Papers

Adjerid I and others, 'Sleights of Privacy: Framing, Disclosures, and the Limits of Transparency' (Proceedings of the Ninth Symposium on Usable Privacy and Security Newcastle, 24-26 July 2013)

Agrawal D, Das S and El Abbadi A, 'Big Data and Cloud Computing: Current State and Future Opportunities' (Proceedings of the 14th International Conference on Extending Database Technology, Uppsala, 21-24 March 2011)
Binns R and others, 'Third Party Tracking in the Mobile Ecosystem' (Proceedings of the 10th ACM Conference on Web Science, Amsterdam, 27-30 May 2018)
Böhme R and Köpsell S, 'Trained to Accept?: A Field Experiment on Consent Dialogs' (Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, 10-15 April 2010)
Cahn A and others, 'An Empirical Study of Web Cookies' (Proceedings of the 25th International Conference on World Wide Web, Montréal, 11-15 April 2016)
Farahat A and Bailey M, 'How Effective is Targeted Advertising?' (Proceedings of the 21st International Conference on World Wide Web, Lyon, 16-20 April 2012)
Thode W, Griesbaum J and Mandl T, '"I would have never allowed it": User Perception of Third-party Tracking and Implications for Display Advertising' (Re:inventing Information Science in the Networked Society Proceedings of the 14th International Symposium on Information Science (ISI 2015), Zadar, 19-21 May 2015)
Weinberg Z and others, 'I Still Know What You Visited Last Summer: Leaking Browsing History via User Interaction and Side Channel Attacks' (2011 IEEE Symposium on Security and Privacy, Berkeley, 22-25 May 2011)
Yan J and others, 'How Much Can Behavioral Targeting Help Online Advertising?' (Proceedings of the 18th International Conference on World Wide Web, Madrid, 20-24 April 2009)

Official Documents

Article 29 Data Protection Working Party, 'Recommendation 1/99 on Invisible and Automatic Processing of Personal Data on the Internet Performed by Software and Hardware' (1999) 5093/98/EN/final WP 17
——, 'Opinion 4/2007 on the concept of personal data' (2007) 01248/07/EN WP 136
——, 'Opinion 1/2009 on the proposals amending Directive 2002/58/EC on privacy and electronic communications (e-Privacy Directive)' (2009) 00350/09/EN WP 159
——, 'Opinion 2/2010 on online behavioural advertising' (2010) 00909/10/EN WP 171
——, 'Opinion 15/2011 on the definition of consent' (2011) 01197/11/EN WP187
——, 'Opinion 01/2012 on the data protection reform proposals' (2012) 00530/12/EN WP 191
——, 'Opinion 04/2012 on Cookie Consent Exemption' (2012) 00879/12/EN WP 194
——, 'Opinion 03/2013 on purpose limitation' (2013) 00569/13/EN WP 203
——, 'Working Document 02/2013 providing guidance on obtaining consent for cookies' (2013) 1676/13/EN WP 208
——, 'Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC' (2014) 844/14/EN WP 217
——, 'Opinion 9/2014 on the application of Directive 2002/58/EC to device fingerprinting' (2014) 14/EN WP 224
——, 'Cookie sweep combined analysis' (2015) 14/EN WP 229
——, 'Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679' (2017) 17/EN WP 251rev.01
——, 'Guidelines on consent under Regulation 2016/679' (2018) 17/EN WP259 rev.01
Commission, 'Proposal for a Directive of the European Parliament and of the Council concerning the processing of personal data and the protection of privacy in the electronic communications sector' (2000) COM/2000/0385 final

——, 'Communication from the Commission to the European Parliament pursuant to the second subparagraph of Article 251 (2) of the EC Treaty concerning the common position of the Council on the adoption of a Directive of the European Parliament and of the Council on processing of personal data and the protection of privacy in the electronic communications sector' (2002) SEC/2002/0124

——, 'Proposal for a Directive of the European Parliament and of the Council amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on consumer protection cooperation' (2007) COM(2007) 698 final

——, 'Amended proposal for a Directive of the European Parliament and of the Council Amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sectors and Regulation (EC) No 2006/2004 on consumer protection cooperation' (2008) COM(2008)723 final

——, 'Communication from the Commission — Guidance on the Commission's enforcement priorities in applying Article 82 of the EC Treaty to abusive exclusionary conduct by dominant undertakings' (2009) OJ C 45/7

——, 'Proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regards to the processing of personal data and on the free movement of such data (General Data Protection Regulation)' (2012) COM(2012) 11 final

——, 'A Digital Single Market Strategy for Europe' (2015) Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions COM(2015) 192 final

——, 'Proposal for a Regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications)' (2017) COM(2017) 10 final

Council, 'Common Position adopted by the Council on 28 January 2002 with a view to the adoption of a Directive of the European Parliament and of the Council concerning the processing of personal data and the protection of privacy in the electronic communications sector' (2002) 15396/2/01 REV 2

——, 'Common position adopted by the Council on 16 February 2009 with a view to the adoption of a Directive of the European Parliament and of the Council amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on consumer protection cooperation' (2009) 16497/1/08 REV 1

——, 'Proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)' (2013) ST 11013 13 INIT

——, 'Proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) - Agreement on key issues of Chapters I-IV' (2013) ST 9398 13 INIT

——, 'Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free

movement of such data (General Data Protection Regulation) - Chapter II, preparation of trilogue' (2015) ST 11245 15 INIT
——, 'Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) - Preparation of a general approach' (2015) ST 9565 2015 INIT
——, 'Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) - Written debriefing of trilogue on 24 November' (2015) ST 14461 15 INIT
——, 'Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation) [first reading] - Analysis of the final compromise text with a view to agreement' (2015) ST 15039 2015 INIT
——, 'Position of the Council at first reading with a view to the adoption of a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL 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) - Statement of the Council's reasons' (2016) ST 5419 2016 REV 1 ADD 1
European Data Protection Supervisor, 'Privacy and competitiveness in the age of big data: The interplay between data protection, competition law and consumer protection in the Digital Economy' (2014) Preliminary Opinion of the European Data Protection Supervisor
——, 'Opinion 3/2018 on online manipulation and personal data' (2018)
European Parliament, 'Second Report on the proposal for a European Parliament and Council directive concerning the processing of personal data and the protection of privacy in the electronic communications sector ' (2001) A5-0374/2001
——, 'Recommendation for Second Reading on the Council common position for adopting a European Parliament and Council directive concerning the processing of personal data and the protection of privacy in the electronic communications sector' (2002) A5-0130/2002
——, 'Report on the proposal for a directive of the European Parliament and of the Council amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on consumer protection cooperation ' (2008) A6-0318/2008
——, 'Recommendation for Second Reading on the Council common position for adopting a directive of the European Parliament and of the Council amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities for the enforcement of consumer protection laws' (2009) A6-0257/2009
——, 'Report on the proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)' (2013) A7-0402/2013
——, 'Report on the proposal for a regulation of the European Parliament and of the Council concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications)' (2017) A8-0324/2017

Federal Trade Commission, 'Self-Regulatory Principles for Online Behavioral Advertising: Tracking, Targeting and Technology' (2009) FTC Staff Report
——, 'Protecting Consumer Privacy in an Era of Rapid Change: A Proposed Framework for Businesses and Policymakers (Preliminary FTC Staff Report)' (2010)
Information Commissioner's Office, 'Determining What Is Personal Data' (2012)
——, 'Guidance on the rules of cookies and similar technologies' (2012) V. 3
——, 'Guide to the Privacy and Electronic Communications Regulations' (2016)
——, 'Consultation: GDPR Consent Guidance' (2017)

Reports

Advertising Association, *Advertising Pays: How Advertising Fuels the UK Economy* (2013) <http://www.adassoc.org.uk/wp-content/uploads/2014/09/Advertising_Pays_Report.pdf> accessed 20 February 2015
——, *Advertising Pays 3: The Value of Advertising to the UK's Culture, Media and Sport* (2015) <http://www.adassoc.org.uk/wp-content/uploads/2015/01/Advertising-Pays-3.pdf> accessed 20 February 2015
Alphabet Inc., *Form 10-K* (2018) <https://abc.xyz/investor/pdf/20171231_alphabet_10K.pdf> accessed 8 June 2018
An M, *Why People Block Ads: And What It Means for Marketers and Advertisers* (2016) <https://cdn2.hubspot.net/hubfs/53/assets/hubspot.com/research/reports/Why_People_Block_Ads.pdf.zip> accessed 16 March 2017
ANA and others, *Programmatic: Seeing Through the Financial Fog - An In-Market Analysis of Programmatic Media at the Transaction Level* (2017) <http://www.ana.net/getfile/25070> accessed 13 March 2018
Article 29 Data Protection Working Party, *Article 29 Data Protection Working Party comments in response to W3C's public consultation on the W3C Last Call Working Draft, 24 April 2014, Tracking Preference Expression (DNT)* (2014) <http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2014/20140606_wp29_ts_standardisation_letter_to_w3c.pdf> accessed 18 January 2018
——, *Article 29 Data Protection Working Party comments in response to W3C's public consultation on the W3C Last Call Working Draft, 14 July 2015, Tracking Compliance and Scope* (2015) <http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2015/20151001__letter_of_the_art_29_wp_w3c_compliance.pdf> accessed 18 January 2018
Barauskas S and Gondard P, *Google : End Of The Online Advertising Bubble* (2016) <https://kalkis-research.com/google-end-of-the-online-advertising-bubble> accessed 27 October 2016
Cisco, *The Zettabyte Era: Trends and Analysis* (2016) <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.pdf> accessed 11 August 2016
Commission, *The EU Data Protection Reform and Big Data Factsheet* (2016) <http://ec.europa.eu/justice/data-protection/files/data-protection-big-data_factsheet_web_en.pdf> accessed 22 February 2017
Dentsu Aegis Network, *Global Ad Spend Forecasts* (2018) <http://www.dentsuaegisnetwork.com/m/en-UK/DentsuAegisAdSpend/DANJune2018.pdf> accessed 8 June 2018
Department for Culture MaS, *Open Letter on the UK Implementation of Article 5(3) of the e-Privacy Directive on Cookies* (2011) <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/77638/cookies_open_letter.pdf> accessed 18 January 2018
European Data Protection Supervisor, *Opinion of the European Data Protection Supervisor on Promoting Trust in the Information Society by Fostering Data Protection and Privacy* (2010)

<https://edps.europa.eu/sites/edp/files/publication/10-03-19_trust_information_society_en.pdf> accessed 18 January 2018

European Parliament, *Data Protection Review: Impact on EU Innovation and Competitiveness* (2012) <http://bookshop.europa.eu/is-bin/INTERSHOP.enfinity/WFS/EU-Bookshop-Site/en_GB/-/EUR/ViewPublication-Start?PublicationKey=BA3112305> accessed 26 March 2017

——, *Party Financing and Referendum Campaigns in EU Member States* (2015) <http://www.europarl.europa.eu/RegData/etudes/STUD/2015/519217/IPOL_STU(2015)519217_EN.pdf> accessed 18 April 2018

Evans D, *The Internet of Everything: How More Relevant and Valuable Connections Will Change the World* (2012) <http://www.lehigh.edu/~inengrit/dropbox/eac1113/Cisco_Internet-of-Everything.pdf> accessed 22 February 2017

Executive Office of the President of the United States, *Big Data and Differential Pricing* (2015) <https://www.whitehouse.gov/sites/default/files/docs/Big_Data_Report_Nonembargo_v2.pdf> accessed 4 October 2016

Facebook Inc., *Form 10-K* (2016) <https://s21.q4cdn.com/399680738/files/doc_financials/annual_reports/2015-Annual-Report.pdf> accessed 22 February 2017

——, *Facebook Annual Report 2016* (2017) <https://s21.q4cdn.com/399680738/files/doc_financials/annual_reports/FB_AR_2016_FINAL.pdf> accessed 1 August 2017

——, *Form 10-K* (2018) <http://d18rn0p25nwr6d.cloudfront.net/CIK-0001326801/c826def3-c1dc-47b9-99d9-76c89d6f8e6d.pdf> accessed 8 June 2018

Filippov S, *Data-Driven Business Models: Powering Startups in the Digital Age* (2014) <www.lisboncouncil.net/component/downloads/?id=1081> accessed 1 November 2016

Google, *Google White Paper: The Arrival of Real-Time Bidding, and What it Means for Media Buyers* (2011) <http://static.googleusercontent.com/media/www.google.com/en//doubleclick/pdfs/Google-White-Paper-The-Arrival-of-Real-Time-Bidding-July-2011.pdf> accessed 18 May 2016

IAB Europe, *Consumers Driving the Digital Uptake: The Economic Value of Online Advertising-Based Services for Consumers* (2010) <http://www.iabeurope.eu/files/7113/7000/0832/white_paper_consumers_driving_the_digital_uptake.pdf> accessed 20 February 2015

——, *Attitudes towards Programmatic Advertising* (2016) <http://www.iabeurope.eu/wp-content/uploads/2016/07/IAB-Europe-Attitudes-towards-Programmatic-Advertising-report_June-2016-v3.pdf> accessed 2 November 2016

IAB Europe and IhS, *AdEx Benchmark 2015: European Online Advertising Expenditure* (2016) <http://www.iabeurope.eu/wp-content/uploads/2016/07/IAB-Europe_AdEx-Benchmark-2015-report_July-2016-V2.pdf> accessed 2 November 2016

——, *European Programmatic Market Sizing 2016* (2017) <https://www.iabeurope.eu/wp-content/uploads/2017/09/IAB-Europe_European-Programmatic-Market-Sizing-2016-report_FINAL-with-appendix.pdf> accessed 13 March 2018

IHS, *Paving the Way: How Online Advertising Enables the Digital Economy of the Future* (2015) <http://www.iabeurope.eu/files/9614/4844/3542/IAB_IHS_Euro_Ad_Macro_FINALpdf.pdf> accessed 28 October 2016

Information Commissioner's Office, *Changes to the Rules on Using Cookies and Similar Technologies for Storing Information* (2011) <https://web.archive.org/web/20110511051726/http:/www.ico.gov.uk/~/media/documents/library/Privacy_and_electronic/Practical_application/advice_on_the_new_cookies_regulations.pdf> accessed 18 January 2018

——, *Guidance on the Rules on Use of Cookies and Similar Technologies* (2011) <https://web.archive.org/web/20111213204833/http://www.ico.gov.uk/news/latest_news/2011/~/media/

documents/library/Privacy_and_electronic/Practical_application/guidance_on_the_new_cookies_regulations.ashx> accessed 18 January 2018

Interactive Advertising Bureau, *Economic Value of the Advertising-Supported Internet Ecosystem* (2012) <http://www.iab.net/media/file/iab_Report_September-24-2012_4clr_v1.pdf> accessed 20 February 2015

Internet Advertising Bureau UK, *A Guide to Online Behavioural Advertising* (2009) <https://www.iabuk.net/sites/default/files/publication-download/OnlineBehaviouralAdvertisingHandbook_5455.pdf> accessed 7 January 2017

——, *BIS Consultation on Implementing the Revised EU Electronic Communications Framework: IAB UK Response* (2010) <https://www.iabuk.net/sites/default/files/IABUKresponsetoBISconsultationonimplementingtherevisedEUElectronicCommunicationsFramework_7427_0.pdf> accessed 18 January 2018

——, *IAB Affiliate Marketing Council Consumer Transparency Framework: A Guide for Publishers / Affiliates* (2013) <https://iabuk.net/sites/default/files/Consumer%20Transparency%20Framework%20V1.2.pdf> accessed 7 March 2017

——, *The Data Deal: How Data Driven Digital Advertising Benefits UK Citizens* (2014) <http://www.iabuk.net/sites/default/files/The%20Data%20Deal%20-%20How%20Data%20Driven%20Digital%20Advertising%20Benefits%20UK%20Citizens.pdf> accessed 20 February 2015

Kristol D and Montulli L, *HTTP State Management Mechanism* (2000) <https://www.ietf.org/rfc/rfc2965.txt> accessed 13 March 2018

Laney D, *3D Data Management: Controlling Data Volume, Velocity and Variety* (2001) <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf> accessed 20 February 2015

LinkedIn Corporation, *Form 10-K* (2016) <https://s21.q4cdn.com/738564050/files/doc_financials/annual/2015/LinkedInAnnualReport_2016.PDF> accessed 20 November 2017

MAGNA, *Global Advertising Forecast* (2017) <https://www.magnaglobal.com/wp-content/uploads/2017/12/121117-MAGNA-Global-Forecast_Winter-Update_Final.pdf> accessed 8 June 2018

McKinsey & Company, *Big Data: The Next Frontier for Innovation, Competition, and Productivity* (2011) <http://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/big-data-the-next-frontier-for-innovation> accessed 13 February 2017

Nielsen, *2016 Nielsen Social Media Report – Social Studies: A Look at the Social Landscape* (2017) <www.nielsen.com/content/dam/corporate/us/en/reports-downloads/2017-reports/2016-nielsen-social-media-report.pdf> accessed 20 November 2017

Office of Fair Trade, *Online Targeting of Advertising and Prices: A Market Study* (2010) <http://webarchive.nationalarchives.gov.uk/20140402142426/http:/www.oft.gov.uk/shared_oft/business_leaflets/659703/OFT1231.pdf> accessed 7 January 2017

Open Data Institute, *Data Sharing and Open Data for Banks: A Report for HM Treasury and Cabinet Office* (2014) <https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/382273/141202_API_Report_FINAL.PDF> accessed 13 March 2018

PubMatic, *Understanding Real-Time Bidding (RTB) From the Publisher Perspective* (2010) <http://pubmaticblog.com/wp-content/uploads/2014/06/Understanding_RTB_Q12010.pdf> accessed 20 March 2015

Purcell K, Brenner J and Rainie L, *Search Engine Use 2012* (2012) <http://www.pewinternet.org/files/old-media//Files/Reports/2012/PIP_Search_Engine_Use_2012.pdf> accessed 1 December 2016

PwC, *Insurance 2020: Turning Change into Opportunity* (2012) <http://www.pwc.com/gx/en/insurance/pdf/insurance-2020-turning-change-into-opportunity.pdf> accessed 22 February 2017

TNS, *2009 Study: Consumer Attitudes About Behavioral Targeting* (2009) <https://dsimg.ubm-us.net/envelope/104162/339732/1249993418527_TRUSTe_TNS_2009_BT_Study_Summary.pdf> accessed 20 March 2017

Tow Center for Digital Journalism, *Post-Industrial Journalism: Adapting to the Present* (2012) <http://towcenter.org/wp-content/uploads/2012/11/TOWCenter-Post_Industrial_Journalism.pdf> accessed 27 March 2017

Twitter Inc., *Form 10-K* (2018) <http://files.shareholder.com/downloads/AMDA-2F526X/6318657145x0x972387/4CA553F3-44F1-48CB-8CB5-F03C6FDE95FE/2017_Annual_Report.pdf> accessed 8 June 2018

U.S. Department of Justice Office of Justice Programs and National Institute of Justice, *Privacy in the Information Age: A Guide for Sharing Crime Maps and Spatial Data* (2001) <https://www.ncjrs.gov/pdffiles1/nij/188739.pdf> accessed 7 January 2017

World Federation of Advertisers, *WFA Guide to Programmatic Media: What Every Advertiser Should Know about Media Markets* (2014) <http://www.wfanet.org/media/programmatic.pdf> accessed 3 November 2016

Zogby Analytics, *Interactive Survey of US Adults* (2013) <http://www.aboutads.info/resource/image/Poll/Zogby_DAA_Poll.pdf> accessed 16 March 2017

——, *Public Opinion Survey on Value of the Ad-supported Internet* (2016) <http://digital.daaoperations.org/sites/digital.daaoperations.org/files/DAA_files/ZogbyAnalyticsConsumerValueStudy2016.pdf> accessed 16 March 2017

Media Coverage

'Google User Data to Be Merged Across All Sites Under Contentious Plan' *The Guardian* (25 January 2012) <https://www.theguardian.com/technology/2012/jan/25/google-merge-user-data-privacy> accessed 25 July 2017

Allen J and Abbruzzese J, 'Cambridge Analytica's Effectiveness Called Into Question Despite Alleged Facebook Data Harvesting' *NBC* (20 March 2018) <https://www.nbcnews.com/politics/politics-news/cambridge-analytica-s-effectiveness-called-question-despite-alleged-facebook-data-n858256> accessed 10 June 2018

Assange J, 'Assange: Google Is Not What It Seems' *Newsweek* (23 October 2014) <http://europe.newsweek.com/assange-google-not-what-it-seems-279447> accessed 7 January 2017

Barbaro M and Zeller Jr. T, 'A Face Is Exposed for AOL Searcher No. 4417749' *The New York Times* (9 August 2006) <http://www.nytimes.com/2006/08/09/technology/09aol.html> accessed 7 January 2016

Betters E, 'New Windows 10 Privacy Dashboard Gives You More Control over Data' *Pocket-Init* (11 January 2017) <http://www.pocket-lint.com/news/139971-new-windows-10-privacy-dashboard-gives-you-more-control-over-data> accessed 25 July 2017

Booth R, 'Inquiry Launched into Targeting of UK Voters Through Social Media ' *The Guardian* (17 May 2017) <https://www.theguardian.com/technology/2017/may/17/inquiry-launched-into-how-uk-parties-target-voters-through-social-media> accessed 25 July 2017

Busvine D, 'Facebook Abused Dominant Position, Says German Watchdog' *Reuters* (19 December 2017) <https://uk.reuters.com/article/us-facebook-germany-dataprotection/facebook-abused-dominant-position-says-german-watchdog-idUKKBN1ED10J> accessed 1 March 2018

Caddy B, 'Google Tracks Everything You Do. Here's How to Delete It' *Wired* (14 August 2017) <http://www.wired.co.uk/article/google-history-search-tracking-data-how-to-delete> accessed 13 March 2018

Cadwalladr C, 'The Great British Brexit Robbery: How Our Democracy Was Hijacked ' *The Guardian* (7 May 2017) <https://www.theguardian.com/technology/2017/may/07/the-great-british-brexit-robbery-hijacked-democracy> accessed 25 July 2017

Cadwalladr C and Graham-Harrison E, 'Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach' *The Guardian* (17 March 2018) <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election> accessed 10 June 2018

Chiel E, 'When Google Is More Powerful than the Government' *Fusion* (11 May 2016) <http://fusion.net/story/301142/google-bans-payday-loan-ads-adwords/> accessed 7 January 2017

Chmielewski D, 'How 'Do Not Track' Ended Up Going Nowhere' *Recode* (4 January 2016) <https://www.recode.net/2016/1/4/11588418/how-do-not-track-ended-up-going-nowhere> accessed 18 January 2018

Constine J, 'Facebook Is Shutting Down Its API for Giving Your Friends' Data to Apps' *TechCrunch* (28 April 2015) <https://techcrunch.com/2015/04/28/facebook-api-shut-down/> accessed 3 May 2018

Daniel E, 'Five Years On, What Has Changed Since the Edward Snowden Scandal?' *Verdict* <https://www.verdict.co.uk/snowden-scandal-five-years-gdpr/> accessed 21 June 2018

Davies H, 'Ted Cruz Using Firm That Harvested Data on Millions of Unwitting Facebook Users' *The Guardian* (11 December 2015) <https://www.theguardian.com/us-news/2015/dec/11/senator-ted-cruz-president-campaign-facebook-user-data> accessed 10 June 2018

Demos T and Seetharaman D, 'Facebook Isn't So Good at Judging Your Credit After All' *The Wall Street Journal* (24 February 2016) <http://www.wsj.com/articles/lenders-drop-plans-to-judge-you-by-your-facebook-friends-1456309801> accessed 7 January 2017

Deutsche Welle, 'Deutsche Post Defends Voter-microtargeting Data Practice' *Deutsche Welle* (2 April 2018) <http://www.dw.com/en/deutsche-post-defends-voter-microtargeting-data-practice/a-43223747> accessed 8 April 2018

Foy H, 'Google Warns Red Tape Threatens European Tech Sector ' *Financial Times* (17 April 2016) <https://www.ft.com/content/97185d98-0225-11e6-9cc4-27926f2b110c> accessed 26 March 2017

Gibbs S, 'Facebook "Tracks All Visitors, Breaching EU Law"' *The Guardian* (31 March 2015) <https://www.theguardian.com/technology/2015/mar/31/facebook-tracks-all-visitors-breaching-eu-law-report> accessed 7 January 2017

Gold J, 'IE Exploit Can Track Mouse Cursor Movements - Even When You're Not in IE' *Techworld* (13 December 2012) <http://www.techworld.com/news/security/adobe-releases-security-updates-for-flash-player-coldfusion-3416178/> accessed 27 June 2017

Goldhill O, 'The Psychology Behind Cambridge Analytica Is Massively Overhyped' *Quartz* (29 March 2018) <https://qz.com/1240331/cambridge-analytica-psychology-the-science-isnt-that-good-at-manipulation/> accessed 10 June 2018

Grassegger H and Krogerus M, 'The Data That Turned the World Upside Down' *Vice* (28 January 2017) <https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win> accessed 20 November 2017

Greenwald G, 'XKeyscore: NSA Tool Collects "Nearly Everything a User Does on the Internet"' *The Guardian* (31 July 2013) <https://www.theguardian.com/world/2013/jul/31/nsa-top-secret-program-online-data> accessed 7 January

Hautala L, 'Facebook Privacy Settings Make You Work to Stop the Data Sharing' *CNET* (22 March 2018) <https://www.cnet.com/news/how-to-stop-sharing-facebook-data-after-cambridge-analytica-mess/> accessed 3 May 2018

Helft M, 'Google Makes a Case That It Isn't So Big' *The New York Times* (28 June 2009) <http://www.nytimes.com/2009/06/29/technology/companies/29google.html> accessed 13 March 2018

Hindman M, 'How Cambridge Analytica's Facebook Targeting Model Really Worked – According to the Person Who Build It' *Independent* (13 April 2018) <https://www.independent.co.uk/life-style/gadgets-and-tech/how-cambridge-analytica-s-facebook-targeting-model-really-worked-according-to-the-person-who-built-a8289901.html> accessed 10 June 2018

Jones R, 'Which? Accuses Admiral of Driver Discrimination' *The Guardian* (16 March 2010) <https://www.theguardian.com/business/2010/mar/16/which-accuses-admiral-insurance-discrimination> accessed 7 January 2017

Kaltheuner F, 'Cambridge Analytica Explained: Data and Elections' *Medium* (13 April 2017) <https://medium.com/privacy-international/cambridge-analytica-explained-data-and-elections-6d4e06549491> accessed 1 March 2018

Kleinman Z, 'Facebook to Update Trending Topics' *BBC* (24 May 2016) <http://www.bbc.co.uk/news/technology-36367216> accessed 7 January 2017

Leibowitz WR, 'Personal Privacy and High Tech: Little Brothers Are Watching You' *The National Law Journal* (7 April 1997)

McGee M, 'As Google Analytics Turns 10, We Ask: How Many Websites Use It?' *Marketing Land* (12 November 2015) <https://marketingland.com/as-google-analytics-turns-10-we-ask-how-many-websites-use-it-151892> accessed 13 March 2018

McGrory R, 'The Data Providers: One Quadrant Chart To Rule Them All' *AdExchanger* (21 February 2013) <https://adexchanger.com/data-driven-thinking/the-data-providers-one-quadrant-chart-to-rule-them-all> accessed 22 February 2017

Naughton J, 'The Rise of Ad-blocking Could Herald the End of the Free Internet' *The Guardian* (27 September 2015) <https://www.theguardian.com/commentisfree/2015/sep/27/ad-blocking-herald-end-of-free-internet-ios9-apple> accessed 3 May 2018

O'Connor P, 'Political Ads Take Targeting to the Next Level' *The Wall Street Journal* (14 July 2014) <http://www.wsj.com/articles/political-ads-take-targeting-to-the-next-level-1405381606> accessed 7 January 2017

Oppenheim M, 'Grindr to Stop Sharing HIV Status of Users with Third-party Companies after Fierce Criticism' *The Independent* (3 April 2018) <https://www.independent.co.uk/life-style/gadgets-and-tech/grindr-hiv-status-users-third-party-companies-stop-sharing-criticism-gay-dating-app-a8286366.html> accessed 3 May 2018

Osborne H and Elgot J, 'MPs Summon Mark Zuckerberg, Saying Facebook Misled Them' *The Guardian* (21 March 2018) <https://www.theguardian.com/uk-news/2018/mar/20/officials-seek-warrant-to-enter-cambridge-analytica-hq> accessed 8 April 2018

Out-Law.com, 'Advertisers Say That New Cookie Law Is Met by Browser Settings' *Out-Lawcom* (24 November 2009) <https://www.out-law.com/page-10550> accessed 18 January 2018

Palmer B, 'Why Is Microsoft Fighting So Hard Over Internet Explorer?' *Slate* (17 December 2009) <http://www.slate.com/articles/news_and_politics/explainer/2009/12/why_is_microsoft_fighting_so_hard_over_internet_explorer.html> accessed 18 January 2018

Picard E, 'The Ethical Issues with 3rd Party Behavioral Tracking' *AdExchanger* (31 October 2011) <https://adexchanger.com/the-debate/3rd-party-behavioral-tracking/> accessed 26 September 2017

Press G, 'A Very Short History of Big Data' *Forbes* (9 May 2013) <https://www.forbes.com/sites/gilpress/2013/05/09/a-very-short-history-of-big-data/> accessed 13 March 2018

——, '12 Big Data Definitions: What's Yours?' *Forbes* (3 September 2014) <http://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours> accessed 22 February 2017

Privacy International, 'Cambridge Analytica Explained: Data and Elections' *Medium* (13 April 2017) <https://medium.com/privacy-international/cambridge-analytica-explained-data-and-elections-6d4e06549491> accessed 10 June 2018

Rankin J, 'Facebook Fined £94m for 'Misleading' EU over WhatsApp Takeover ' *The Guardian* (18 May 2017)

Rosenberg M, Confessore N and Cadwalladr C, 'How Trump Consultants Exploited the Facebook Data of Millions' *The New York Times* (17 March 2018) <https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html> accessed 10 June 2018

Ruddick G, 'Facebook Forces Admiral to Pull Plan to Price Car Insurance Based on Posts' *The Guardian* (2 November 2016) <https://www.theguardian.com/money/2016/nov/02/facebook-admiral-car-insurance-privacy-data> accessed 30 May 2018

Shearman S, 'Europe's "chilling" Data Reforms Will Deter a UK Facebook or Twitter' *Campaign* (2 May 2012) <http://www.campaignlive.co.uk/article/europes-chilling-data-reforms-will-deter-uk-facebook-twitter/1129829> accessed 24 March 2017

Sloane G, 'Facebook Lends Trending Hand to Brands' *Adweek* (17 January 2014) <http://www.adweek.com/digital/facebook-lends-trending-hand-brands-155060> accessed 22 February 2017

Sweney M, 'Online Paid-content Market Poses Threat to Traditional Advertising' *The Guardian* (1 November 2012) <https://www.theguardian.com/media/2012/nov/01/online-paid-contect-rise-8-billion-pounds> accessed 25 July 2017

The Guardian, 'Google User Data to Be Merged Across All Sites Under Contentious Plan' *The Guardian* (25 January 2012) <https://www.theguardian.com/technology/2012/jan/25/google-merge-user-data-privacy> accessed 18 March 2018

——, 'The Guardian View on Data Protection: Informed Consent Needed' *The Guardian* (19 March 2018) <https://www.theguardian.com/commentisfree/2018/mar/19/the-guardian-view-on-data-protection-informed-consent-needed> accessed 8 April 2018

Wagner K, 'Pinterest Expects to Make More Than $500 Million in Revenue This Year' *Recode* (21 March 2017) <https://www.recode.net/2017/3/21/14991260/pinterest-advertising-revenue-500-million-growth-ipo> accessed 20 November 2017

Wakefield J, 'Cambridge Analytica: Can Targeted Online Ads Really Change a Voter's Behaviour?' *BBC* (30 March 2018) <https://www.bbc.co.uk/news/technology-43489408> accessed 10 June 2018

Weintraub S, 'Google Continues Buying Spree, Picks Up Invite Media' *Fortune* (2 June 2010) <http://fortune.com/2010/06/02/google-continues-buying-spree-picks-up-invite-media/> accessed 13 March 2018

Wheatley M, 'Gartner Warns Big Data's Bubble May Burst as Enterprises Cut Investment' *SiliconANGLE* (4 October 2016) <https://siliconangle.com/blog/2016/10/04/gartner-warns-big-datas-bubble-may-burst-as-enterprises-plot-investments-elsewhere/> accessed 6 June 2018

Wheeler B, 'GCHQ Could "Grab" UK Shopping Data, Committee Told' *BBC* (10 December 2015) <http://www.bbc.co.uk/news/uk-politics-35060064> accessed 7 January 2017
Wong JC, 'Facebook's Privacy Practices are Under Investigation, FTC Confirms' *The Guardian* (26 March 2018) <https://www.theguardian.com/technology/2018/mar/26/facebook-data-privacy-cambridge-analytica-investigation-ftc-latest> accessed 8 April 2018
Woodie A, 'Why Gartner Dropped Big Data Off the Hype Curve' *Datanami* (26 August 2015) <https://www.datanami.com/2015/08/26/why-gartner-dropped-big-data-off-the-hype-curve/> accessed 6 June 2018

Webpages

Advertising Association and Warc, 'Shift to Digital Will Push UK Advertising to a Record £20bn+ in 2015' (2014) <http://expenditurereport.warc.com/FreeContent/Q4_2013.pdf> accessed 26 March 2017
——, 'Adspend Growth Forecasts Hold Strong Despite Brexit Vote' (2017) <http://expenditurereport.warc.com/FreeContent/Q3_2016.pdf> accessed 26 March 2017
Albright J, 'Cambridge Analytica: the Geotargeting and Emotional Data Mining Scripts' (2017) <https://medium.com/tow-center/cambridge-analytica-the-geotargeting-and-emotional-data-mining-scripts-bcc3c428d77f> accessed 20 November 2017
Article 29 Data Protection Working Party, 'Letter of 16 December 2016' (2016) <https://ec.europa.eu/newsroom/document.cfm?doc_id=40927> accessed 21 February 2018
——, 'Letter of 24 October 2017' (2017) <https://ec.europa.eu/newsroom/just/document.cfm?doc_id=47964> accessed 21 February 2018
CNIL, 'Data Transfer from WHATSAPP to FACEBOOK: CNIL Publicly Serves Formal Notice for Lack of Legal Basis' (2017) <https://www.cnil.fr/en/data-transfer-whatsapp-facebook-cnil-publicly-serves-formal-notice-lack-legal-basis> accessed 21 February 2018
Datanyze, 'Web Analytics Market Share Report' <https://www.datanyze.com/market-share/web-analytics> accessed 13 March 2018
Dutcher J, 'What Is Big Data?' (2014) <https://datascience.berkeley.edu/what-is-big-data> accessed 22 February 2017
EDAA, 'Your Online Choices' <http://www.youronlinechoices.com/uk/about-behavioural-advertising> accessed 26 September 2017
European Data Coalition, 'GDPR Redlines – The Missing Link Between GDPR and DSM' (2015) <http://europeandatacoalition.eu/wp-content/uploads/2015/06/GDPR-redlines-the-missing-link-between-the-dsm-and-the-gdpr.pdf> accessed 26 March 2017
European Data Protection Supervisor, 'Data Protection' <https://edps.europa.eu/data-protection/data-protection_en> accessed 10 June 2018
European Interactive Digital Advertising Alliance, 'YourOnlineChoices.eu - About' <http://www.youronlinechoices.com/uk/about-behavioural-advertising> accessed 7 March 2017
Facebook, 'Facebook Elections' <https://www.facebook.com/business/a/politics-industry> accessed 27 June 2016
——, 'How Can I Download My Information from Facebook?' <https://www.facebook.com/help/212802592074644> accessed 7 January 2017
——, 'What Is Online Interest-based Advertising from Facebook, and How Can I Control Whether I See Online Interest-based Ads?' <https://www.facebook.com/help/164968693837950> accessed 13 March 2018
Gartner, 'Gartner Hype Cycle' <http://www.gartner.com/technology/research/methodologies/hype-cycle.jsp> accessed 13 March 2018

Glanert M, 'Emerging Trends in Online Advertising' (2010) <http://behavioraltargeting.biz/emerging-trends-in-online-advertising/> accessed 3 March 2018
Google, 'About Ads Based on Websites That You've Visited' <https://support.google.com/ads/answer/1697735> accessed 13 March 2018
——, 'About Advertising Features' <https://support.google.com/analytics/answer/3450482> accessed 13 March 2018
——, 'About Google Ads' <https://support.google.com/ads/answer/1634057> accessed 13 March 2018
——, 'About Personalised advertising' <https://support.google.com/adsense/answer/113771> accessed 13 March 2018
——, 'Advertising ID' <https://support.google.com/googleplay/android-developer/answer/6048248> accessed 13 March 2018
——, 'big data - Explore - Google Trends' <https://trends.google.com/trends/explore?date=all&q=big%20data> accessed 22 February 2017
——, 'Big Data, Machine Learning, Blockchain, IoT - Explore - Google Trends' <https://trends.google.co.uk/trends/explore?date=today%205-y&geo=US&q=Big%20Data,Machine%20Learning,Blockchain,IoT> accessed 6 June 2018
——, 'Cambridge Analytica - Explore - Google Trends' <https://trends.google.co.uk/trends/explore?q=Cambridge%20Analytica> accessed 10 June 2018
——, 'Cookie Matching | Real-Time Bidding Protocol | Google Developers' <https://developers.google.com/ad-exchange/rtb/cookie-guide> accessed 22 February 2017
——, 'Data Collection | How Google Uses Your Personal Information' <https://privacy.google.com/intl/en-GB/your-data.html> accessed 7 March 2017
——, 'Delete Directions and Places from Your History' <https://support.google.com/maps/answer/3137804> accessed 13 March 2018
——, 'Download Your Data' <https://support.google.com/accounts/answer/3024190> accessed 7 January 2017
——, 'Get Started | Real-Time Bidding Protocol | Google Developers' <https://developers.google.com/ad-exchange/rtb/start> accessed 22 February 2017
——, 'Google Scholar' <https://scholar.google.co.uk/scholar?q=%22big+data%22&as_ylo=2017&as_yhi=2017> accessed 13 March 2018
——, 'How Does Facebook Decide Which Ads to Show Me and How Can I Control the Ads I See?' <https://www.facebook.com/help/562973647153813> accessed 13 March 2018
——, 'How Google Infers Interest and Demographic Categories' <https://support.google.com/adsense/answer/140378> accessed 13 March 2018
——, 'How Google Uses Data When You Use Our Partners' Sites or Apps' <https://www.google.com/policies/privacy/partners/> accessed 13 March 2018
——, 'Our History in Depth' <https://www.google.co.uk/about/company/history> accessed 21 December 2016
——, 'Real-Time Bidding Protocol Buffer v.138' <https://developers.google.com/ad-exchange/rtb/downloads/realtime-bidding-proto> accessed 13 March 2018
——, 'Topics Used for Personalised Ads' <https://support.google.com/ads/answer/2842480> accessed 13 March 2018
——, 'Win the Moments That Win Elections' <https://www.google.com/ads/elections/> accessed 7 January 2017
——, 'Google Acquires Applied Semantics' (2003) <http://googlepress.blogspot.co.uk/2004/04/google-acquires-applied-semantics.html> accessed 13 March 2018
——, 'Here comes Measure Map' (2006) <https://googleblog.blogspot.co.uk/2006/02/here-comes-measure-map.html> accessed 22 February 2017

——, 'The DoubleClick Ad Exchange: Growing the Display Advertising Pie for Everyone' (2009) <https://googleblog.blogspot.co.uk/2009/09/doubleclick-ad-exchange-growing-display.html> accessed 22 February 2017
——, 'We've Officially Acquired AdMob!' (2010) <https://googleblog.blogspot.co.uk/2010/05/weve-officially-acquired-admob.html> accessed 27 May 2010
——, 'Policy Requirements for Google Analytics Advertising Features' (2016) <https://support.google.com/analytics/answer/2700409> accessed 13 March 2018
——, 'Privacy Policy' (2017) <https://static.googleusercontent.com/media/www.google.co.uk/en/uk/intl/en_uk/policies/privacy/google_privacy_policy_en_uk.pdf> accessed 26 September 2017
IAB Europe, 'Data Protection Deal Is a Setback for Europe's Digital Economy' (2015) <https://www.iabeurope.eu/policy/data-protection-deal-is-a-setback-for-europes-digital-economy/> accessed 24 March 2017
IETF, 'Internet Protocol' (1981) <https://tools.ietf.org/html/rfc791> accessed 7 January 2017
Industry Coalition for Data Protection, 'Europe's New Data Rules Take a Wrong Turn' (2015) <http://epceurope.eu/wp-content/uploads/2015/12/ICDP-Press-Release_Final.pdf> accessed 24 March 2017
Interactive Advertising Bureau, 'Telecom Package: Publishers and Online Marketers Welcome New Provisions on Cookies' (2009) <http://www.iab.it/iab-news/telecom-package-publishers-and-online-marketers-welcome-new-provisions-on-cookies/> accessed 18 January 2018
Internet Advertising Bureau UK, 'Industry Concerned over EU Data Protection Developments' (2013) <https://iabuk.net/news/industry-concerned-over-eu-data-protection-developments> accessed 24 March 2017
——, 'EU Governments Approve Privacy Text That Threatens Online Business' (2015) <https://iabuk.net/about/press/archive/eu-governments-approve-privacy-text-that-threatens-online-business> accessed 24 March 2017
Library of Congress, 'Congress.gov' <https://www.congress.gov/advanced-search/legislation?enterTerms=%22Do+Not+Track%22&search=search> accessed 18 January 2018
LUMA, 'Display LUMAscape' <https://www.lumapartners.com/luma-institute/lumascapes/display-ad-tech-lumascape/> accessed 13 March 2018
Mayer J, 'Tracking the Trackers: Microsoft Advertising' (2011) <http://cyberlaw.stanford.edu/blog/2011/08/tracking-trackers-microsoft-advertising> accessed 13 March 2018
Mozilla, 'About Lightbeam' <https://www.mozilla.org/en-GB/lightbeam/about/> accessed 7 January 2017
Narayanan A, 'Online Price Discrimination: Conspicuous by Its Absence' (2013) <https://33bits.org/2013/01/08/online-price-discrimination-conspicuous-by-its-absence/> accessed 7 January 2017
Netflix, 'Netflix Error N8011' <https://help.netflix.com/en/node/71> accessed 7 January 2017
Network Advertising Initiative, 'US Internet Users Understand and Value Interest-Based Advertising, According to DAA Survey' (2013) <http://www.networkadvertising.org/blog/us-internet-users-understand-and-value-interest-based-advertising-according-daa-survey> accessed 19 March 2017
Ranker, 'The Top Advertising Data Exchanges and Aggregators' <http://www.ranker.com/list/advertising-data-exchanges-and-aggregators/online-ad-network-lists> accessed 22 February 2017
ShareThis, 'Privacy Notice' <http://www.sharethis.com/privacy> accessed 22 February 2017
The Hamburg Commissioner for Data Protection and Freedom of Information, 'Administrative Order Against the Mass Synchronisation of Data Between Facebook and

WhatsApp' (2016) <https://www.datenschutz-hamburg.de/fileadmin/user_upload/documents/Press_Release_2016-09-27_Adminstrative_Order_Facebook_WhatsApp.pdf> accessed 21 February 2018
University of Bristol, 'Linkage to Routine Health and Social Records' <http://www.bristol.ac.uk/alspac/researchers/our-data/linkage/> accessed 13 March 2018
——, 'Researchers' <http://www.bristol.ac.uk/alspac/researchers> accessed 13 March 2018
——, 'Insomnia More Common in Teens Whose Mums Dad Postnatal Depression' (2016) <http://www.bristol.ac.uk/alspac/news/2016/depression-and-insomnia.html> accessed 13 March 2018
W3C, 'Tracking Protection Working Group' <https://www.w3.org/2011/tracking-protection/> accessed 18 January 2018
——, 'Header Field Definitions' (1999) <http://www.w3.org/Protocols/rfc2616/rfc2616-sec14.html> accessed 7 January 2017
——, 'Tracking Compliance and Scope: W3C Candidate Recommendation 26 April 2016' (2016) <https://www.w3.org/TR/tracking-compliance/> accessed 18 January 2018
——, 'Tracking Preference Expression (DNT): W3C Candidate Recommendation 19 October 2017' (2017) <https://www.w3.org/TR/tracking-dnt/> accessed 18 January 2018
Wang X and others, 'United States Patent: Generating User Profiles' (2013) <https://patentimages.storage.googleapis.com/87/fb/2c/bf0a5e5a68f605/US8352319.pdf> accessed 13 March 2018
Wojtkiewicz R, 'A Day in the Light: A 24-Hour Experiment with Mozilla's Lightbeam' (2013) <https://thebottomline.as.ucsb.edu/2013/11/a-day-in-the-light-a-24-hour-experiment-with-mozillas-lightbeam> accessed 7 January 2017