Policy Brief

Data Policies: Towards Citizen-Centric Regulation

Arne Hintz

IT for Change | January 2020

This report was produced as part of the research project ‘Policy frameworks for digital platforms - Moving from openness to inclusion’. The project seeks to explore and articulate institutional-legal arrangements that are adequate to a future economy that best serves the ideas of development justice. This initiative is led by IT for Change, India, and supported by the International Development Research Centre (IDRC), Canada.

Author

Arne Hintz is Senior Lecturer at the School of Journalism, Media, and Culture at Cardiff University and co-Director of Data Justice Lab.

Research coordination team

Principal Investigator: Anita Gurumurthy
Co-investigators: Deepti Bharthur, Nandini Chami
Editorial Support: Visvak, Vinay Narayan
Design: Purnima Singh

© IT for Change 2020

Licensed under a Creative Commons License Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)

Data Policies: Towards Citizen-Centric Regulation

Policy Brief

January 2020

Arne Hintz

Data Justice Lab

Cardiff University

Page intentionally left blank

Summary

The datafication of social life is leading to a profound transformation in the manner in which society is ordered, decisions are made and citizens are governed. The rules and norms that regulate the collection and use of data are, therefore, crucial cornerstones of this emerging datafied society.

Recent policy reforms suggest that this regulatory environment is developing in two seemingly opposite directions. On the one hand, we are witnessing increased collection and sharing of personal data by state agencies. However, on the other hand, data protection and citizens' control over their data is also being enhanced. Data collection and dataveillance have become pervasive, but the need to empower citizens to control the data that characterizes them is gaining increasing recognition. This is happening through normative frameworks such as data ethics and legislation like the EU's General Data Protection Regulation (GDPR).

User empowerment can be an important part of data policies and the GDPR offers many useful starting points, such as the right to explanation, data portability, and improved consent rules. However, an approach that focuses on individual user responsibility has limitations and is insufficient to address the challenges of a datafied society. Normative frameworks and user empowerment must be complemented by rigorous legislative restrictions that regulate the exploitation of people's data.

These laws should include robust regulation of the key processes of datafication, such as profiling, data sharing, and automated decision-making. They should also restrict the collection of data by both commercial and state actors as collection itself can cause significant harmful effects. Moreover, data policies need to incorporate regulations that deal with derived and inferred data, complementing the narrower notion of personal data.

The policy debate around data will need to move beyond individual approaches to data control—such as individual privacy—and consider collective concepts of data. Further, in order to maintain and advance democracy in a datafied society, opportunities for civic participation in the development and rolling-out of data-based decision-making systems need to be established.

Data collection and analysis has allowed commercial and state institutions to predict and change human behaviour, and to sort, categorize, and assess citizens

1. Introduction

The datafication of social life has led to a profound transformation in the ways in which society is ordered, decisions are made, and citizens are governed. The capacity to analyze big data has created opportunities "to extract new insights or create new forms of value" (Mayer-Schönberger & Cukier, 2013, p. 8).

Datafication has come to define contemporary life: our society has been termed a datafied society (Hintz et al., 2018) and the current political-economic system has been described as surveillance capitalism (Zuboff, 2019). Data collection and analysis has allowed commercial and state institutions to predict and change human behaviour, and to sort, categorize, and assess citizens. This significantly affects the roles of citizens and the protection and understanding of their civic rights. The rules and norms that regulate the collection and use of data are therefore crucial cornerstones of emerging societal formations.

This has become a particularly prominent concern with the proliferation of social media platforms, cloud services, and the so-called sharing economy, whose core business model is the collection, analysis, and monetization of user data. Platforms are “data mines” (Andrejevic, 2012) from where personal data is systematically extracted, processed, and combined with additional datasets in order to create detailed profiles of people that are valuable to business interests. Their increasingly fundamental role in contemporary society has been conceptualized using terms such as ‘platform society’ (van Dijk et al., 2018) and ‘platform capitalism’ (Srnicek, 2016). Despite their increasing importance, they have largely operated in a policy vacuum and most of their activities remain unregulated. The far-reaching social, political, and economic consequences of the actions of platforms point to a pressing need for policy development in this space.

This policy brief reviews current trends in the regulation of data collection and analysis with a focus on platforms. In particular, it interrogates emerging regulatory frameworks that shape, constrain, or advance citizens’ control over personal (and related) data. In doing so, it points to areas of necessary intervention to address citizen needs and concerns.

In order to ground the argument in specific current policy developments and controversies, this policy brief focuses on the national and regional jurisdictions of the UK and the EU. Thus, it is set against a backdrop of multiple pronounced controversies involving data collection and analysis—from the Snowden revelations and the Cambridge Analytica-Facebook scandal, to attempts to enhance citizen control over data through the GDPR. Further, it is situated in a contradictory regulatory environment, with the collection and use of personal data being enhanced by state surveillance legislation and, at the same time, restricted by data protection legislation—and with new policy norms exploring the critical intersections of innovation, security, citizen rights, and user autonomy.

The policy brief is based on empirical research conducted in 2018 by members of the Data Justice Lab at Cardiff University, UK. This research project was part of a collaborative global research project on platform policies led by IT for Change, India.

2. Data: What is the Problem?

Described as “the new oil” (The Economist, 2017), data analysis has facilitated a vast new business sector that aims to predict consumption patterns based on a variety of social, cultural, health and other information (McCann et al., 2018). It is also transforming public services, which are increasingly allocated based on data analytics about claimants, leading to automated welfare eligibility systems and the use of predictive risk models in, for example, child protective services and the health sector (Eubanks, 2018). In criminal justice systems and border control, risk assessment tools are used to produce risk scores on defendants and migrants to estimate their future conduct (Angwin et al., 2016; Metcalfe & Dencik, 2019). Recent debates on the use of data analytics in the UK have involved sectors such as policing, criminal justice, housing, and child welfare, as well as the deployment of data analytics provided by credit rating agencies for public services (Dencik et al., 2018).

This has led to increasing concerns about potential “data harms” (Redden & Brand, 2018). The pervasive monitoring and surveillance of citizens through the collection and analysis of their data traces has been discussed extensively since the Snowden revelations on mass surveillance by intelligence and security agencies, including the UK’s Government Communications Headquarters (Lyon, 2015). Further, many critics have highlighted the possibility that data analytics will entrench existing forms of discrimination (Gangadharan et al., 2015). Moreover, the black box nature of big data processes—i.e. the lack of understanding regarding how an algorithm processes data and arrives at certain results—raises questions of transparency and accountability regarding the working and consequences of datafication. This poses a significant problem as it renders people incapable of understanding, investigating, and challenging the processes by which society is increasingly organized (Pasquale, 2015) and thus casts doubt on the agency of supposedly active and informed digital citizens (Hintz et al., 2018).

Many scholars have critically interrogated the premise of big data and its associated algorithmic processes. They have criticized ideas of data having a value-neutral, impartial, and objective character as "carefully crafted fictions" (Kitchin, 2014) and have noted that data is always constructed based on the goals, interests, and cultures of institutions and individuals. The representation of reality by data and, more specifically, the relationship between people and the data that is collected about them, is thus not self-evident. Data analytics, moreover, provides a reduced lens on society (Berry, 2011) and shapes the reality it measures by focusing on specific objects and on certain methods of knowing and understanding social life (Boyd & Crawford, 2012). Scholars have also raised concerns regarding the "operative logic of pre-emption" (Massumi, 2015) inherent in data-based governance that challenges existing practices and understandings of the democratic process (Andrejevic, 2017) by focusing on managing the consequences of social ills rather than seeking to understand their underlying causes.

3. Predominant Regulatory Approaches

3.1 Self-Regulation and Co-Regulation

In the digital economy, tentative interpretations of user consent have formed the core of self-regulatory data regimes. Such regimes typically require platforms and apps to seek acceptance from users for the ways in which companies track their browsing habits and use their data. For example, the EU Directive on Privacy and Electronic Communications, which was issued in 2002 and amended in 2009, required explicit consent from websites visitors for the installation of cookies that could identify, track, and profile them. However, under this model of regulation, platforms and services typically require users to agree to comprehensive collection of their data if they wish to partake in digital life. Thus, this model places the burden of privacy protection on the individual and "merely legitimises the extraction of personal data from unwitting data subjects" (Edwards & Veale, 2017, p. 49).

3.2 Data Protection and Privacy Legislation

Data protection legislation, such as the UK's Data Protection Act, has controlled the collection, use, and sharing of personal data by companies and provided limitations for the same. Such national rules are embedded in regional and international policy such as the European Convention for the Protection of Human Rights and Fundamental Freedoms, which was incorporated into UK law in the Human Rights Act, 1998. Article 8 of the Convention guarantees everyone the "right to respect for his private and family life, his home and his correspondence" (Council of Europe, 1950).

3.3 Surveillance Legislation and Exemptions

Data protection laws often have significant exemptions that allow the state to collect and share data for the protection of national security and the prevention or detection of crime. The UK Regulation of Investigatory Powers Act from 2000, which was amended by the Data Retention and Investigatory Powers Act, 2014, allowed a Secretary of State to authorize the interception of wide-ranging and vaguely defined types of traffic in bulk in addition to the communications of specific individuals. Similar powers have been included in other relevant laws, such as the Telecommunications Act, 1984 and the Wireless Telegraphy Act, 2006 (Hintz & Brown, 2017).

3.4 Normative Approaches

Non-binding policy norms can guide the development of legislation and provide a useful environment for a debate on what should and should not be done. Such norms may include statements and declarations by institutions such as the UN, the Council of Europe, and national policy advisory organizations. Pressure by civil society groups and other stakeholders may also lead to changes in norms and affect the range and legitimacy of available regulatory options.

4. Regulatory Trends and Examples

4.1 Data Collection by the State: The UK Investigatory Powers Act

Despite high-profile scandals such as the Snowden revelations, many states have expanded their legal frameworks for data collection. Typically, this has been justified using security considerations in light of terrorist threats. The UK Investigatory Powers Act from 2016 is a particularly far-reaching example. It is a comprehensive piece of legislation intended to combine the previously fragmented rules for communications interception and data collection by state agencies. The Act addresses a wide range of surveillance practices, including targeted and bulk interception, mandatory communications data retention by platforms, the collection of internet connection records (i.e., people's web browsing habits), and computer network exploitation (i.e., hacking by state agencies into servers and devices). While the law opens up many surveillance measures that were traditionally secret to public scrutiny and oversight, it also confirms, legalizes, and expands existing practices of state-based data collection and analysis (Hintz & Brown, 2017).

Regulatory reform has allowed for the expansion of data collection and for the increased use of data by the state

The Digital Economy Act from 2017 mandates data collection by private sector entities by requiring certain platforms to establish age verification procedures for their users and age-related content filtering. It also facilitates data sharing between government departments as well as between government and private companies without citizens' knowledge, thus transferring control of personal data away from the citizen. As the examples of these two laws demonstrate, regulatory reform has allowed for the expansion of data collection and for the increased use of data by the state.

4.2 Data Protection and Restrictions Against Commercial Use: The EU General Data Protection Regulation

Meanwhile, in apparent contradiction to the trend of extended data collection, data protection rules have been strengthened in some jurisdictions, leading to increased regulation of the data-related activities of the internet industry and commercial platforms. The most prominent case is the GDPR from 2018—a comprehensive regulatory framework that limits the use and sharing of personal data by companies and provides citizens with some control in the context of several new challenges that have emerged with datafication. For example, the law mandates purpose limitation of data collection and processing, limits the processing of sensitive personal data (personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs), prohibits decisions based solely on automated processing, requires impact assessments for potentially harmful data uses, and mandates data protection by design.

The GDPR expands and refines requirements for user consent and gives users the option to withhold or withdraw consent (and still use the service). It includes rules regarding the right of access to personal data and data portability, which make it easier for users to switch from one provider to another. It also makes automated and algorithmic decision-making more transparent, assigning citizens the right to demand an explanation and to challenge the outcomes of algorithmic decisions. The regulation thus puts a strong focus on expanding user control, which may address power imbalances between the state and platforms, on the one hand, and between the state and citizens, on the other. However, the GDPR places the onus for user protection and the burden of proof—in the case of the right to explanation and challenge—on the data subject, i.e. the individual citizen. Enhancements to the need for consent and the right to explanation may create a "transparency fallacy" as "individual data subjects are not empowered to make use of the kind of algorithmic explanations they are likely to be offered" as they lack "the necessary expertise to meaningfully make use of these individual rights" (Veale & Edwards, 2017, p. 66-67). Further, the GDPR rules have limitations to their scope and thus, their effectiveness. They do not, for example, include inferred data that is generated from observation of people's online activities, which has become the main source of profiling (Wachter, et al., 2017). The GDPR, therefore, offers significant improvements for citizens' data rights, yet with shortcomings and limitations.

5. Norms: Data Autonomy, Data Ethics, and the Informed User

The goals of the GDPR are increasingly reflected in national normative statements and institutional developments. The UK government's Digital Charter from 2018 includes the provision that "personal data should be respected and used appropriately". The UK's new Centre for Data Ethics and Innovation addresses concerns regarding the collection and analysis of personal data by, for example, reviewing potential biases of algorithmic decision-making. Data ethics has become a popular normative framework internationally for data use in both public institutions and the private sector. This demonstrates the growing recognition of the various effects of data on citizens and of the need to protect citizens and enhance their capabilities in a datafied environment.

However, just like parts of the GDPR, the normative approaches of data autonomy and data ethics focus on the role of the informed user and thereby place responsibility on the citizen (or, in part, the company), not the government. Thus, they individualize the regulatory framework of datafication and may distract from the need to develop adequate laws and regulations. Further, ethical data use does not prevent extensive levels of data collection. Data ethics frameworks are often discussed in connection with the alleged need for data collection for innovation (in a business context) or security (in a state context). Thus, data ethics may legitimize extended levels of collection and use of personal data. While such norms encourage user control over data, they are situated within a perceived need to advance the digital economy overall.

6. Recommendations

Measures that ensure an active and informed role for citizens in a datafied society are crucial to data policies. These include rules laid out in the GDPR, such as the right of access to personal data and to data portability, and the right to explanation that allows citizens to understand how their data is used. These measures should be complemented by a robust right to challenge the outcomes of algorithmic decisions. However, over-reliance on individual responsibility will not be an effective approach to addressing the challenges of a datafied society. Rigorous legislative restrictions on the exploitation of people's data must include the purpose limitation of data collection and analysis, limits to profiling, to data sharing between public and private institutions, and to automated decision-making.

The collection of data by both commercial and state actors is an integral part of the potentially harmful effects of datafication. Robust policy mechanisms therefore need to restrict the collection, as well as the analysis, use, and sharing, or data. Data ethics can be a useful complementary framework but cannot replace legislation. Without being accompanied by robust regulation, it risks transforming the protection of citizen rights into a self-guided act by public and private sector entities that is either voluntary or negotiated between those stakeholders.

Data policies (including the rights to access and data portability, and limits to profiling) need to extend to inferred and derived data, i.e. the wide range of behavioral, locational, click stream, and other data that is becoming more valuable than the rather limited personal data that data subjects knowingly provide.

As data denotes a relation to others and the individual's place within a broader collective, data policies need to expand beyond the limits of individual approaches. Innovative approaches to a collective understanding of data have emerged, for example, through the concept of indigenous data sovereignty whereby indigenous communities have formulated programs that advocate the right to maintain, control, protect, and develop data that is collected about them (Kukutai & Taylor, 2016).

In order to maintain and advance democracy in a datafied society, mechanisms need to be evolved that will ensure civic participation in the development and rolling-out of data-based decision-making systems

Data localization can be part of such a collective control over data and has been proposed as a means to address the geopolitical debates that have arisen around the concentration of data processing capabilities in the US. However, by virtue of being a national approach, it risks potentially advancing national data collection and surveillance strategies, thereby enhancing governmental rather than citizen control. Localization policies at the municipal level, connected with other strategies of decentralisation, may hold more promise.

As datafication is a trend of broader societal transformation, data-specific policies (including data protection legislation and data ethics norms) need to be situated within the context of broader societal processes. Most importantly, in order to maintain and advance democracy in a datafied society, mechanisms need to be evolved that will ensure civic participation in the development and rolling-out of data-based decision-making systems.

Changes in policy require normative and discursive changes. Citizens' rights and control, as a policy goal, competes with security (or rather, a specific understanding of national security) and innovation (i.e., allowing data use with limited restrictions) as a leading benchmark. The protection of citizens and the enhancement of their control over data that concerns them have become more prominent goals recently but will have to assert their place against other frames.

References

Andrejevic, M. (2012). Exploitation in the data mine. In C. Fuchs, K. Boersma, A. Albrechtslund & M. Sandoval (Eds.), *Internet and Surveillance: The Challenges of Web 2.0 and Social Media* (71–88). Abingdon: Routledge.

Andrejevic, M. (2017). To pre-empt a thief. *International Journal of Communication*, 11, 879–96.

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine bias. *Pro Publica*, Retrieved from https://www.propublica.org

Berry, D. (2011). The computational turn: Thinking about the digital humanities. *Sussex Research Online*, 12, 1–22. Retrieved from https://sro.sussex.ac.uk/id/eprint/49813/1/berry_2011-the_computational_turn-_thinking_about_the_digital_humanities.pdf

boyd, d. & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662–679.

Council of Europe. (1950). *European Convention of Human Rights*. Retrieved from https://www.echr.coe.int/Documents/Convention_ENG.pdf

Dencik, L., Hintz, A., Redden, J. & Warne, H. (2018). Data Scores as governance: Investigating uses of citizen scoring in public services. Retrieved from Data Justice Lab website https://datajustice.files.wordpress.com/2018/12/data-scores-as-governance-project-report2.pdf

The Economist (2017, May 6). Fuel of the Future: Data is giving rise to a new economy. *The Economist*. Retrieved from: https://www.economist.com/news/briefing/21721634-how-it-shaping-up-data-giving-rise-new-economy

Edwards, L. & Veale, M. (2017). Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for. *Duke Law and Technology Review*, 16(1), 18–84.

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. New York: St Martin's Press.

Gangadharan, S.P., Eubanks, V. & Barocas, S. (Eds.) (2015). *Data and Discrimination: Collected Essays*. New America: Washington DC.

Hintz, A. & Brown, I. (2017). Enabling digital citizenship? The reshaping of surveillance policy after Snowden. *International Journal of Communication*, 11, 782–801.

Hintz, A., Dencik, L. & Wahl-Jorgensen, K. (2018). *Digital citizenship in a datafied society*. Cambridge: Polity Press.

Kitchin, R. (2014). *The data revolution*. London: Sage.

Lyon, D. (2015) *Surveillance after Snowden*. Cambridge: Polity.

Massumi, B. (2015). *Ontopower: War, powers, and the state of perception*. Durham, NC: Duke University Press.

Mayer-Schönberger, V. & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work and think*. New York: John Murray.

McCann, D., Hall, M. & Warin, R. (2018). Controlled by calculations?: Power and accountability in the digital economy. Retrieved from https://neweconomics.org/2018/06/controlled-by-calculations

Metcalfe, P. & Dencik, L. (2019). The politics of big borders: Data (in)justice and the governance of refugees. *First Monday*, 24(4) Retrieved from https://firstmonday.org/ojs/index.php/fm/article/view/9934/7749

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Cambridge, MA: Harvard University Press.

Redden, J. & Brand, J. (2018). Data Harm Record. Retrieved from https://datajusticelab.org/data-harm-record/

Srnicek, N. (2016). *Platform Capitalism*. Cambridge: Polity.

van Dijk, J., Poell, T., & de Waal, M. (2018). *The Platform Society*. Oxford: Oxford University Press.

Veale, M. & Edwards, L. (2017). Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling. *Computer Law & Security Review*, 34(2), 398–404.

Wachter, S., Mittelstadt, B. & Floridi, L. (2017). Why a right to explanation of automated decision-making does not exist in the General Data Protection Regulation. *International Data Privacy Law*, 7(2), 76–99.

Zuboff, S. (2019). *The Age of Surveillance Capitalism*. London: Profile Books.