The scored consumer: privacy and big data

Peissl, Walter; Krieger-Lamina, Jaro

Veröffentlichungsversion / Published Version
Konferenzbeitrag / conference paper

Empfohlene Zitierung / Suggested Citation:
Peissl, W., & Krieger-Lamina, J. (2017). The scored consumer: privacy and big data. In C. Bala, & W. Schuldzinski (Eds.), *The 21st Century Consumer: Vulnerable, Responsible, Transparent? ; Proceedings of the International Conference on Consumer Research (ICCR) 2016* (pp. 101-112). Düsseldorf: Verbraucherzentrale Nordrhein-Westfalen e.V. https://doi.org/10.15501/978-3-86336-918-7_9

The scored consumer

Privacy and big data

Dr Walter Peissl (Austria) | Institute of Technology Assessment of the Austrian Academy of Sciences

Jaro Krieger-Lamina (Austria) | Institute of Technology Assessment of the Austrian Academy of Sciences

DOI 10.15501/978-3-86336-918-7_9

1 Introduction

The 'scored consumer' is just one example of many generic developments in and around big data, entailing potential conflicts with the fundamental right to privacy. In this paper, we will discuss the relation between elements of the big data debate and the issue of privacy, and we will exemplify the challenges arising by looking into the details of credit scoring. The current challenges originate predominantly from socio-technical developments during recent decades, which will continue to unfold over the coming years. Three of these socio-technical and techno-economic-driven developments are digitisation, big data and the Internet of things. These developments are very relevant to all of us, as they are supposed to change—and have already changed—our lives profoundly.

First and foremost, we see a broad wave of digitisation. Digitisation as a socio-technical phenomenon is not new in itself. Since the 1970s we have witnessed an ever-increasing diffusion of digital applications. From the professional use of information and communication technologies (ICT) in large organisations, offices and factories alone, initially using central computers, we saw the widespread adoption of personal computers in the late 1980s and early 1990s and the digital doll that is currently in our children's nurseries. Nowadays, we use digital equipment in almost every area or dimension of our daily lives, and we are constantly producing data. The inherent and technically necessary production of data by digital systems represents a fundamental difference compared to older analogous systems. The data produced is readily available and it seems an obvious choice to use it. Thus, measurement and control have become central paradigms. In contrast to analogous systems, data does not vanish in digital systems, having been produced. For instance, in old analogous telephone systems, after an actual call, no traces of the call could be found in the system. If, for example, the law enforcement authorities wanted to interfere or eavesdrop, they had to catch the actual conversation 'on the fly'. In general, data such as caller ID, time and duration of call are produced and saved in digital systems and would have to be actively erased. Most are stored for an indefinite period of time and can be used later.

Today, most people use digital equipment and services in their everyday lives, such as smartphones, debit and credit card payments, different self-tracking devices and wearables, which may give rise to the phenomenon of self-tracking and self-optimisation, known under the header of the 'quantified self' (Selke 2016). There have also been digital developments in caregiving, such as Paro the robot. In the domain of health care, we are increasingly obliged to use digital applications for administrative use (e.g. patient cards), as well as for individual monitoring. A field of ample use of ICT is the security sector. Security issues have dominated political debate for more than fifteen years. The so-called securitisation of societies (Buzan, Wæver and de Wilde 1998) smoothes the way for heavy (digital) surveillance all over the world. Digitisation promises to deliver efficiency gains and to bring about new products and services. At the same time, problems from power and market imbalances are becoming more acute.

A second phenomenon is big data—a hype and a buzzword. Big data is an overarching term for new developments in ICT to exploit the huge and unstructured data that is generated every second by different systems. Experts often refer to the 5 Vs of big data: volume, velocity, variety, veracity and value (Demchenko, Ngo, and Membrey 2013). Volume refers to the vast amount of data generated every second. Velocity refers to the speed at which new data is generated and the speed at which data moves around. Variety refers to the different types of data that can be used now. And veracity refers to the messiness or (lack of) trustworthiness of the data. These 4 Vs basically refer to increasing technical potential. The fifth V directly points at a very important trigger: 'But all the volumes of fast-moving data of different variety and veracity have to be turned into value! This is why value is the one V of big data that matters the most. …Value refers to our ability to turn our data into value. It is important that businesses make a case for any attempt to collect and leverage big data' (Marr 2015). From a consumers' point of view, we have to ask who will be paying for the profits of the companies using this data, and what are the mechanisms of turning data into gold.

Another vision that seems to be slowly materialising is the Internet of things: 'IoT refers to the networked interconnection of everyday objects, which are often equipped with ubiquitous intelligence. IoT will increase the ubiquity of the Internet by integrating every object for interaction via embedded systems,

which leads to a highly distributed network of devices communicating with human beings as well as other devices' (Xia et al. 2012). This means that almost everything will become interconnected; every single object will have an identifying number, and these objects will constitute (or be equipped with) embedded systems. The 'smart' components will enable independent communication and decision-making by the systems and all interconnected objects will be potential tracking devices.

The developments from digitisation, big data and the Internet of things may lead, on the one hand, to the 'automation' of society (Helbing 2015). On the other hand, we do have strong legal reasons to govern innovation. One of them, although often attacked and endangered, is the fundamental right to privacy.

2 Privacy—a fundamental right and a prerequisite of democratic societies

There are good reasons for the existence of fundamental legal barriers to the abundant use of all the data generated. One of them is privacy—a fundamental right and a prerequisite of a democratic society (see: Peissl 2014; Pauer-Studer 2003; Wiederin 2003; Bennett 2003). But what is privacy? In short, privacy goes back to a famous article by Warren and Brandeis (1890), where they invented and argued for a 'right to be let alone'. Interestingly, the case was based on the application of a new technology that was already in place. It was the time when personal photography was becoming available, street photography was evolving and pictures were being published in the newspapers. Second, there was a very influential court ruling (BVerfGE 1983) by the German Bundesverfassungsgericht (Federal Constitutional Court) on informational self-determination, which means that every individual should have the right to control who knows what about him- or herself. The Bundesverfassungsgericht interlinked control over the information flow with the free will of autonomous citizens as the basis for democratic societies. Hence, the discussion about privacy is not

only about individual and fundamental rights. It is also about the philosophy of autonomy and freedom, and thus about the pillars of liberal democracy (Rössler 2001; Pauer-Studer 2003). Following Rössler (2001, 25), privacy may be defined as the possibility to control the access of others to oneself. It can be split into three dimensions: first, the dimension of private decisions and actions ('dezisionale Privatheit'), which means that the individual retains control over his/her own decisions and is (relatively) free of heteronomy. The second dimension is called local privacy ('lokale Privatheit'), which means that the individual has control over rooms and spaces such as his or her own home. Information privacy ('informationelle Privatheit') is the third dimension and relates to the knowledge and control of information about oneself. Transgressing this 'privacy paradigm' (Bennett and Raab 2003), we can observe a tendency towards a 'surveillance society' (David Murakami Wood (ed.) et al. 2006) that entails other threats to individuals, such as categorisation and social sorting (Lyon 2003).

The new technological means mentioned earlier create an extreme challenge for information privacy, as well as for the sphere of private decisions and actions. Meanwhile, even local privacy is under pressure: the consumer's voice 'controls' the smart TV in the living room, the smart doll in the nursery, as well as smart speaker-like devices all around the apartment, which means they have to have an 'open ear' to the consumer's wishes. In particular, 'smart TVs' that react to gestures and voice commands and 'smart speakers' such as the Amazon Echo or the Google Home exemplify the strong push towards greater convenience and comfort at the cost of personal data. All the virtual assistants, such as Google Now, Apple's Siri, Amazon's Alexa and Microsoft's Cortana, are intruding on our privacy. Too many consumers are not aware of the implications and therefore happily trade their privacy for a much-valued 'modern' way of living. This goes hand-in-hand with immense societal developments induced and intensified by social media. Billions of people worldwide use social media, sharing details of their lives and those of others, such as their children. We have observed a tremendous impact on the understanding of individual privacy.

3 Potential impacts of consumer mass surveillance

Highly influential political developments accompany these socio-technical changes: in particular, security measures and the political debate over the 'war on terror' have held sway for more than fifteen years and are now paving the way for massive surveillance. Is this the end of privacy (Enserink and Chin 2015)? Are we living in the post-privacy era (Heller 2011)? Or do we have good reason to try to live up to the existing rules while helping people to maintain their private lives? In the following section, we will give a short overview of the potential impacts of the mass surveillance of consumers.

By knowing their customers, enterprises want to improve their goods and services in order to enhance turnover and profit. Therefore, the predictability of consumer behaviour is a major aim. Allegedly, the more data a system has at its disposition, the more it will be able to predict what will happen. As soon as their behaviour becomes predictable, consumers turn into economically exploitable entities and individuals into manageable risk factors. This is a dramatic change in the way we think about human beings.

Another challenge is the process of social sorting (Lyon 2003), which means that people are put into categories based on their virtual 'data double'. In a sense, the virtual data double invades, mimics and eventually determines the real lives of people.

A third issue is mainstreaming: as soon as we know that there is surveillance around, this reduces our privacy, and a loss of diversity results—in other words, there is a strong mainstreaming effect. Under surveillance (and under manipulation through unequal hierarchical communication) people no longer use their full autonomy and decide what they think others might expect from them. In this way, with decisions prepared and guided by systems, we actually do not know what accompanies our daily lives. It may serve the short-term interests of enterprises, which can optimise logistics and raise profits. In the long term, however, this could be dangerous for our societies. Restricting var-

iation and avoiding what some would deem deviation may do away with any 'driving momentum' in societal, cultural and economic terms. Only diversity in behaviour—even if ambivalent—carries the potential for change. Deviant behaviour is the driver of progress. So there could be a long-term impact on our societies' ability to innovate (see: Peissl 2003).

4 Credit scoring—a challenge for consumers

In the context of big data, privacy and consumer mass surveillance, credit scoring is a perfect example of the mechanisms described. Almost nobody in Austria knows anything about credit scoring; not even the mere existence of such a system is publicly known (Rothmann, Sterbik-Lamina, and Peissl 2014). At the same time, we can observe an increasing number of complaints over financial services filed at consumer-protection organisations regarding the potential discrimination of loan-seekers (VKI and AK 2013). This accompanies the excessive growth of large data collections in the private sector and a grey market for potentially illegal data. Scoring bears hallmarks of surveillance, because many different aspects of consumers' private lives are being monitored over time. Basically, scoring systems disregard privacy principles such as the limitation principle and the purpose principle (Rothmann, Sterbik-Lamina, and Peissl 2014).

From a historical perspective, scoring procedures initially served to protect creditors and became established in the business-to-business sector only. Very early on, instruments were blacklists of defaulters, which were published regularly. Relatively early on, specialised companies established scoring or blacklist services and, in addition, expanded into more advanced services such as debt collection. A focus on consumers was not noticeable until the 1960s. In the early phases, only personal knowledge of a few criteria such as occupation, salary and known assets served to assess the creditworthiness of consumers.

Scoring today is based on hard facts such as income, loans, open debts and other banking data. Interestingly, ever more soft facts such as socio-economic data, including address, forenames and surnames, family status, social media profiles, as well as smartphone usage, are fed into the systems. Additionally, sector-specific blacklists influence the individual's score. The score—mostly unknown to the customer—determines the availability and pricing of services such as interest rates of loans, mobile phone contracts and payment options in online shopping platforms. If your score is not good enough, you may have prepaid options only, whereas your neighbour from across the street (beyond the 'red line') also has the option to pay by credit card or even after delivery.

As already mentioned, digitisation and minimisation will spread wearables and tracking devices everywhere. Thus, measuring and quantifying all dimensions of life will increase. In the Internet of things data from all the different sources will be used to intensify the interlinking of different data sets. This will lead to more specific profiles. Hence, scoring will be increasingly based on data doubles rather than facts or 'real' behaviour. The more that economic value becomes decisive for assessing customers, the more social sorting (Lyon 2003) and categorisation may become predominant.

As scores become increasingly important for consumers, it may be a relevant new task for them to actually manage their score. In this respect, much more transparency is needed. Consumers must have the chance to know their score to be able to act accordingly. Consumers have to be enabled to maintain their right of access, of correction and of erasure of their data, as stated in several national and international rules, directives and guidelines for data protection (Galetta and de Hert 2014, Österreichischer Nationalrat 1999). Transparency is also needed with regard to the algorithms applied in scoring systems. Nevertheless, it is questionable whether we want to live in a society in which citizens must relocate to pay lower interest rates for their loans.

As scoring services get cheaper (because of digitisation) and also increasingly available for SMEs, we are already observing an imbalance and disproportionality with regard to business volumes. Even in contracts with low or almost no risk, or very little money involved, scoring systems are used. Hence, the customers' rights to privacy are neglected without due cause.

5 Conclusions

Observing socio-technical innovations such as digitisation, big data and the Internet of things, we anticipate the further growth of scoring systems and their use by companies in different sectors. Departing from the concepts of equality and fairness, there is a need for a new and more comprehensive regulation of these services. With regard to future developments, a specific scoring law would also have to be considered in Austria.

Foremost, the lack of transparency and awareness should be tackled. As long as consumers do not know that they are subject to scoring, they do not have a chance to react accordingly or to request the correction of wrong data. Transparency also includes the knowledge of the algorithms of scoring systems in order to be able to know what kind of data or behaviour might give rise to a certain score. The various assessment and calculation methods must be not only disclosed but also examined by independent supervisory authorities.

The data limitation principle should be applied and the use of certain data restricted. Only directly relevant data on creditworthiness should be used, such as income situation, assets, loans and expected expenditure. In this context, reference must also be made to the purpose principle of data protection: not all the available data on a person may be used for credit scoring. Certain areas of life and particular types of data should be explicitly excluded.

A threshold regarding the contracted financial risk should be installed, preventing scoring for bagatelle transactions. This also implies that services and products that form part of the basic livelihood in Austria, including renting flats, should be exempted from comprehensive credit assessments. Credit scoring should only be permitted for a very limited extent of business transactions involving material goods that can be procured in the event of a default.

Credit rating data may not be available to all employees of scoring companies at all times. Access to the recorded, archived and analysed personal (creditworthiness) data has to be kept to a minimum and recorded in a revision-proof manner. The control of compliance with the provisions should, in the first in-

stance, be entrusted to a company data protection officer, who should be installed in all companies that use data scoring.

In addition, there are strikingly few empirical investigations from the social sciences on the subject of credit scoring and the credit assessment of natural persons. There are no representative findings on the experiences of, and implications for, the consumers concerned. As in other technology fields, the statistical models and algorithms governing the scoring procedures are (unknowingly) used to incorporate the values of the clients and developers. How this affects the results of the scoring calculations requires further scientific investigation.

Finally, the cultural or philosophical question remains about whether we really want to quantify ourselves to squeeze the multidimensional and rich aspects of our lives into different data sets prepared by scoring schemes that apply largely unknown algorithms.

References

Bennett, Colin J. 2003. 'Information privacy and "Datenschutz": Global assumptions and international governance.' In *Privacy: Ein Grundrecht mit Ablaufdatum? Interdisziplinäre Beiträge zur Grundrechtsdebatte*, edited by Walter Peissl, 61–81. Wien: Verlag der Österreichische Akademie der Wissenschaften.

Bennett, Colin J., and Charles D. Raab. 2003. *The governance of privacy*. Aldershot, Hampshire GB: Ashgate.

Buzan, Barry, Ole Wæver, and Jaap de Wilde. 1998. *Security – A new framework for analysis*. Boulder: Lynne Rienner Publisher.

BVerfGE. 1983. BVerfGE 65, 1 – Volkszählung, Urteil des Ersten Senats vom 15. Dezember 1983 auf die mündliche Verhandlung vom 18. und 19. Oktober 1983 – 1 BvR 209, 269, 362, 420, 440, 484/83 in den Verfahren über die Verfassungsbeschwerden, edited by BVerfGE.

Demchenko, Yuri, Canh Ngo, and Peter Membrey. 2013. 'Architecture framework and components for the big data ecosystem, draft version 0.2.' In *System and Network Engineering*, edited by Universiteit van Amsterdam, and System and Network Engineering Group.

Enserink, Martin, and Gilbert Chin. 2015. 'The end of privacy.' *Science* 347 (6221): 490–491. doi: 10.1126/science.347.6221.490.

Galetta, Antonella, and Paul de Hert. 2014. 'A European perspective on data protection and access rights.' In *Increasing Resilience in Surveillance Societies (IRISS) D5: Exercising democratic rights under surveillance regimes*.

Helbing, Dirk. 2015. *The automation of society is next – How to survive the digital revolution.* North Charleston, SC: Createspace.

Heller, Christian. 2011. *Post-Privacy: Prima leben ohne Privatsphäre*. München: Beck.

Lyon, David (ed.). 2003. *Surveillance as social sorting: Privacy, risk and digital discrimination.* London: Routledge.

Marr, Bernard. 2015. 'Why only one of the 5 Vs of big data really matters.' *IBM Big Data & Analytics Hub*. Accessed 6 March 2017.

Murakami Wood, David (ed.), Kirstie Ball, David Lyon, Clive Norris, and Charles Raab. 2006. *A report on the surveillance society*, edited by Information Commissioner. London.

Österreichischer Nationalrat. 1999. Bundesgesetz über den Schutz personenbezogener Daten (Datenschutzgesetz 2000 – DSG 2000).

Pauer-Studer, Herlinde. 2003. 'Privatheit: Ein ambivalenter aber unverzichtbarer Wert.' In *Privacy: Ein Grundrecht mit Ablaufdatum? Interdisziplinäre Beiträge zur Grundrechtsdebatte*, edited by Walter Peissl, 9–22. Wien: Verlag der Österreichische Akademie der Wissenschaften.

Peissl, Walter. 2003. 'Surveillance and security – a dodgy relationship.' *Journal of Contingencies and Crisis Management* 11 (1):19–24.

––– 2014. 'Grundpfeiler freier Existenz – Privacy: by design, by default und warum überhaupt?' *ÖKZ* 55 (6–7): 20–21.

Rössler, Beate. 2001. *Der Wert des Privaten*. Frankfurt am Main: Suhrkamp.

Rothmann, Robert, Jaro Sterbik-Lamina, and Walter Peissl. 2014. *Credit Scoring in Österreich*. Wien: AK Wien.

Selke, Stefan. 2016. *Lifelogging: Digitale Selbstvermessung und Lebensprotokollierung zwischen disruptiver Technologie und kulturellem Wandel.* Wiesbaden: Springer VS.

VKI (Verein für Konsumenteninformation), and AK (Arbeitnehmerkammer). 2013. *Bericht zur Lage der KonsumentInnen 2011/2012*. Wien: Im Auftrag des Bundesministeriums für Arbeit, Soziales und Konsumentenschutz

site/attachments/8/0/6/CH3582/CMS1448291551119/bericht_zur_lage_2011__2012.pdf.

Warren, Samuel D., and Louis D. Brandeis. 1890. 'The right to privacy.' *Harvard Law Review* IV (5): 193–220.

Wiederin, Ewald. 2003. 'Der grundrechtliche Schutz der Privatsphäre: Eine Entwicklungsgeschichte.' In *Privacy: Ein Grundrecht mit Ablaufdatum? Interdisziplinäre Beiträge zur Grundrechtsdebatte*, edited by Walter Peissl, 23–50. Wien: Verlag der Österreichische Akademie der Wissenschaften.

Xia, Feng, Laurence T. Yang, Lizhe Wang, and Alexey Vinel. 2012. 'Internet of things.' *International Journal of Communication Systems* 25 (1): 101–102. doi: 10.1002/dac.2417.