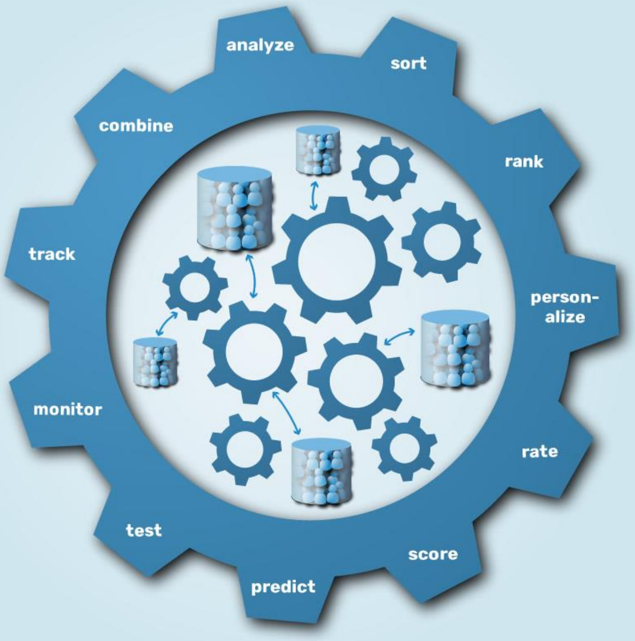
Wolfie Christl

HOW COMPANIES USE PERSONAL DATA AGAINST PEOPLE

Automated Disadvantage, Personalized Persuasion, and the Societal Ramifications of the Commercial Use of Personal Information



WORKING PAPER BY CRACKED LABS

Vienna, October 2017

Author: Wolfie Christl

Contributors: Katharina Kopp, Patrick Urs Riechert

Wolfie Christl

How Companies Use Personal Data Against People

Automated disadvantage, personalized persuasion, and the societal ramifications of the commercial use of personal information.
Working paper by Cracked Labs, Vienna, October 2017

Author: Wolfie Christl
Contributors: Katharina Kopp, Patrick Urs Riechert
Cover illustration: Pascale Osterwalder

Every effort has been made to ensure the accuracy of the texts in this report. The author and the publisher accept no liability in the case of eventual errors.

Cracked Lab – Institute for Critical Digital Culture
Annagasse 8/1/8, 1010 Vienna, Austria
http://crackedlabs.org

Wolfie Christl
is a digital rights activist, technologist, researcher, writer and educator, based in Vienna
http://twitter.com/WolfieChristl
http://wolfie.crackedlabs.org

The production of this report was supported by the Open Society Foundations.

Table of Contents

Abstract

Today, companies aggregate, trade, and utilize personal information at unprecedented levels. Their unilateral and extensive access to data about the characteristics, behaviors, and lives of billions allows them to constantly monitor, follow, judge, sort, rate, and rank people as they see fit. Our previous report documented the massive scale and scope of today's networks of digital tracking and profiling. It investigated relevant industries, business models, platforms, services, devices, technologies, and data flows, focusing on their implications for people – whether as individuals, consumers, or citizens – and society at large.

This working paper examines how the corporate use of personal information can affect individuals, groups of people, and society at large, particularly in the context of automated decisions, personalization and data-driven persuasion. After briefly reviewing our previous research's findings and key developments in recent years, this paper explores their potential to be used against people in detail.

Systems that make decisions about people based on their data produce substantial adverse effects that can massively limit their choices, opportunities, and life-chances. These systems are largely opaque, nontransparent, arbitrary, biased, unfair, and unaccountable – even in areas such as credit rating that have long been regulated in some way. Through data-driven personalization, companies and other institutions can easily utilize information asymmetries in order to exploit personal weaknesses with calculated efficiency. Personalized persuasion strategies provide the means to effectively influence behavior at scale. As companies increasingly and unilaterally shape the networked environments and experiences that underlie and determine everyday life, manipulative, misleading, deceptive, or even coercive strategies can be automated and customized down to the individual level.

Based on the examination of business practices and their implications we conclude that, in their current state, today's commercial networks of digital tracking and profiling show a massive potential to limit personal agency, autonomy, and human dignity. This not only deeply affects individuals, but also society at large. By improving the ability to exclude or precisely target already disadvantaged groups, current corporate practices utilizing personal information tend toward disproportionally affecting these groups and therefore increase social and economic inequality. Especially when combined with influencing strategies derived from neuroeconomics and behavioral economics, data-driven persuasion undermines the concept of rational choice and thus the basic foundation of market economy. When used in political campaigns or in other efforts to shape public policy, it may undermine democracy at large.

While this working paper does not directly offer solutions, it examines, documents, structures, and contextualizes today's commercial personal data industries and their implications; further research will build on this basis. Hopefully, it will also encourage and contribute to further work by others.

Introduction

Not too long ago, the scale and depth of personal information in the hands of commercial entities was quite limited and rather easy to oversee. Credit bureaus, direct marketing firms, and businesses selling products and services to consumers started to collect, manage, and exchange data on people decades ago. That is not to say that people being numbered, rated, and ranked didn't have consequences for many in the past; however, earlier consumer databases were isolated, updated slowly, and captured only a fraction of a typical person's life. Fast-forward to the year 2017 and the situation has changed dramatically. Since the rise of social networks, smartphones, and online advertising, a wide range of companies has started to monitor, track, and follow people across virtually all aspects of their lives. Today, the behaviors, movements, social relationships, interests, weaknesses, and most private moments of billions are constantly recorded, evaluated, and analyzed in real-time.[1]

When surfing the web, hidden pieces embedded of software transmit information about the websites visited, navigation patterns, and sometimes even keystrokes, scrolls and mouse movements to hundreds of third-party companies. Similarly, when carrying a smartphone, rich information about the user's everyday life not only flows to Google, Apple, and a variety of app providers, but also to a significant number of third-party companies, again based on hidden software embedded by app providers. Such information may include a person's contacts, information about real-time app usage and movements, as well as data from all kinds of sensors recording motion, audio, video, and more. Furthermore, as a rapidly increasing number of devices connects to the internet – from wearables, e-readers, TVs, game consoles, toys, baby monitors, printers, and voice-controlled speakers to thermostats, smoke alarms, energy meters, door locks, and vehicles –personal data collection threatens to become ubiquitous and totalizing. Already now, though, individuals can see only the tip of the data and profiling iceberg. Most of it occurs in the background and remains opaque; as a result, most consumers, as well as civil society, journalists, and policymakers, barely grasp the full extent and forms of corporate digital tracking and profiling.[2]

The large-scale and widely unrestrained commercial exploitation of personal data raises major concerns about the future of autonomy, equality, human dignity, and democracy. Our previous report[3] published in June 2017 examined and documented the massive scale and scope of how companies collect, disclose, trade and utilize personal information about individuals. It investigated relevant industries, business models, platforms, services, devices, technologies, and

[1] See Christl, Wolfie (2017): Corporate Surveillance in Everyday Life. How Companies Collect, Combine, Analyze, Trade, and Use Personal Data on Billions. Reporty by Cracked Labs, June 2017. Available at: http://crackedlabs.org/en/corporate-surveillance

[2] Ibid.

[3] Ibid.

data flows, focusing on their implications for people – whether as individuals, consumers, or citizens – and society at large.

This working paper further explores and examines how the corporate **use** of the collected personal information can affect individuals, groups of people, and society, in particular in the context of two partially overlapping areas of concern: automated decisions and data-driven persuasion. Systems that make decisions about people based on their data can massively affect their choices, opportunities, and life-chances. Personalization can be easily abused to exploit personal weaknesses, to persuade people to act in certain ways, and to influence behavior at scale. After briefly reviewing our previous research's finding regarding today's personal data industries, networks of corporate surveillance, and key developments in recent years, this paper explores their potential to be used against people in detail.

According to the upcoming EU General Data Protection Regulation's definition, personal data is "any information relating to an identified or identifiable natural person"[4]. This includes data where identifying information has been replaced by pseudonyms[5] such as numbers or obscure codes. Individuals and society are affected in several ways by the digital processing of personal data. Generally, the rise of the Internet and social media may have changed the ways how people deal with information about their lives and privacy in many cases. Today, some people decide to put details of their private lives online and make them public or semi-public.[6] Others want to have control about what the public, and their friends, family members, neighbors, coworkers, or perhaps certain people from their past can see online.[7] The implications and challenges of the changing ways in which people handle digital information about them at a personal level are diverse[8], but not subject to the considerations of this paper.

Facebook, for instance, has addressed these issues a lot. In the early years, the platform pushed users towards making more and more information about them publicly accessible by default.[9] In recent years, however, the company mostly stopped doing so and has respectably improved the ways users can control their privacy on Facebook at an interpersonal level.[10] Nevertheless,

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

[5] Pseudonymisation is, according to the EU GDPR, the "processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person;"

[6] https://www.instagram.com/

[7] https://www.facebook.com/help/325807937506242/

[8] See e.g. Dumortier, Franck (2009): Facebook and Risks of "De-contextualization" of Information. In: Monograph "5th Internet, Law and Politics Congress. The Pros and Cons of Social Networks", Universitat Oberta de Catalunya. Available at: https://works.bepress.com/franck_dumortier/1/

[9] See: http://mattmckeon.com/facebook-privacy/ [16.09.2017]

[10] Constine, Josh (2014): Facebook Stops Irresponsibly Defaulting Privacy Of New Users' Posts To "Public", Changes To "Friends". TechCrunch, May 22, 2014. Available at: https://techcrunch.com/2014/05/22/sometimes-less-open-is-more/

what users cannot adjust with Facebook's privacy settings is how the platform *itself* takes advantage of the rich digital profiles it stores about users. The latter is what this paper is about. It examines how powerful commercial institutions utilize and exploit digital personal information about individuals who are in a less powerful position,[11] as well the consequences of the resulting power and information asymmetries.

The unilateral and extensive access to data about the characteristics, behaviors, and lives of billions allows companies to constantly evaluate, judge, sort, rank, and single out individuals at unprecedented scale.[12] Furthermore, this also allows them to extract knowledge in order to develop better analysis and data mining technologies, especially in the field of machine learning.[13] While these practices of knowledge extraction create a "new kind of digital divide" between the "Big Data rich" and the "Big Data poor"[14], and thus also raise concerns about information and power asymmetries,[15] this paper mostly focuses on situations where the corporate processing of personal data directly relates to individuals. Of course, the collection of large amounts of personal information is often used for both purposes today. Moreover, extracting knowledge from *big personal data* in a way that is not directly aligned to individuals at first can help improve the analysis, assessment, and classification of individuals based on easily observable personal data later. In this way, both issues are overlapping and related; however, this examination emphasizes the implications of corporate practices utilizing access to personal information in a way that directly relates to – and affects – individuals.

[11] For an overview of the relationship between powerful parties and data subjects see e.g. Rhoen, M. (2016): Beyond consent: improving data protection through consumer protection law. Internet Policy Review, 5(1). Available at: https://policyreview.info/articles/analysis/beyond-consent-improving-data-protection-through-consumer-protection-law

[12] See Christl (2017)

[13] See e.g. Calo, Ryan (2017): Artificial Intelligence Policy: A Roadmap (August 8, 2017), p. 19-21. Available at: https://ssrn.com/abstract=3015350

[14] boyd danah; Crawford, Kate (2012): Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon, Information, Communication & Society 15:5, p. 674. Available at: http://www.tandfonline.com/doi/abs/10.1080/.VB8Tz_l_uCk

[15] Ryan Calo (2017) refers to these practices of knowledge extraction from big data as the „data parity problem" and as a "key policy challenge", see: Ryan Calo

1. (De)centralized mass dataveillance

In 2014, the US Federal Trade Commission examined nine consumer data brokers and found that these companies "gather massive amounts of data, from online and offline sources, and combine them into profiles about each of us".[16] They do so "largely without consumers' knowledge", and, because data brokers also provide each other with data, it would be "virtually impossible for a consumer to determine" how they obtained it.[17] Since then, the situation has become even worse. In recent years, pre-existing practices of commercial consumer data collection have rapidly evolved into pervasive networks of digital tracking and profiling. Today, a vast landscape of partially interconnected databases has emerged that consists not only of large players such as Facebook and Google but also of thousands of other companies from various industries that collect, analyze, share, trade, and utilize data on billions of people.[18]

While this goes on behind the scenes, consumers are left in the dark. They tend not to be aware of what personal information about them and their behavior is collected, nor how this data is processed, with whom it is shared or sold, which conclusions can be drawn from it, and which decisions base on such data. One reason for this certainly lies in the high levels of complexity and abstraction at play. Perhaps more importantly, though, companies make no effort to improve transparency or understanding; on the contrary, they inform consumers incompletely, inaccurately, or not at all, often employing ambiguous, misleading, and obfuscating language. Whether in user interfaces or in contracts, the disclosures that do exist – such as privacy policies and terms of service – are difficult to understand, obscure, and use hypothetical language. Moreover, companies often systematically trick consumers into data contracts. As soon as privacy advocates, consumer rights organizations, regulators, scholars, and journalists ask for more information, companies decline to answer, arguing that their data practices constitute trade secrets and must therefore be protected and kept secret.[19]

Dataveillance. With the scale and depth of today's corporate data collection a reality has materialized that has long been examined under the frame of "surveillance", which the Canadian sociologist and surveillance studies scholar David Lyon defines as the "focused, systematic and routine attention to personal details for purposes of influence, management, protection or direction".[20] Already back in 1988, Roger Clarke coined the notion of "dataveillance" as the "systematic monitoring of people's actions or communications through the application of information technology".[21] He made a distinction between **personal dataveillance**, which concerns

[16] US Federal Trade Commission (2014): Data Brokers. A Call for Transparency and Accountability. May 2014, p. C-3

[17] Ibid., p 46

[18] Christl, Wolfie and Sarah Spiekermann (2016): Networks of Control. A Report on Corporate Surveillance, Digital Tracking, Big Data & Privacy. Facultas, Vienna 2016, p. 7. Available at: http://crackedlabs.org/en/networksofcontrol

[19] Christl and Spiekermann (2016), p. 121-123

[20] Lyon, David (2007): Surveillance Studies: An Overview. Cambridge: Polity Press, p. 14

[21] Roger Clarke (1988): Information Technology and Dataveillance. Commun. ACM 31, 5 (May 1988), 498-512. Available at: http://www.rogerclarke.com/DV/CACM88.html

itself “with identified individuals about whom some kind of concern or suspicion has arisen”, and **mass dataveillance**, which does so “with groups of people and involves a generalized suspicion that some (as yet unidentified) members of the group may be of interest”.[22] While the former investigates transactions and individuals that appear to be exceptional, the latter routinely and automatically monitors and screens all transactions of large groups of people. According to Clarke’s analysis, mass dataveillance bases on profiling and statistical techniques and aims to regulate and control a group’s behavior at scale. It does so by detecting exceptions from pre-defined norms in order to eventually single out, identify, and address individuals.[23]

Social sorting. Decades later, personal dataveillance perhaps corresponds to an insurance investigator manually examining the social connections of a person suspicious of claims fraud with a military-grade data mining tool.[24] In many cases, though, such investigation of exceptions has been automated, too. Today, both personal and mass dataveillance have become the norm and part of everyday life. When today’s social media platforms, credit reporting agencies, consumer data brokers, banks, insurers, telecom companies, loyalty program providers, device providers, and online advertising firms constantly monitor and profile billions of people, they are typically not interested in single natural persons. Rather, they engage in a practice that Lyon, building on the work of Oscar Gandy[25], refers to as “social sorting”; this describes how “personal and group data” are used to constantly “classify people and populations according to varying criteria, to determine who should be targeted for special treatment, suspicion, eligibility, inclusion, access”.[26] As differently classified groups of people are treated differently, this permanent sorting based on data is discriminatory per se and affects the choices and life-chances of individuals.

Identification. While corporate data collection has, for the most part, little interest in specific natural persons, identification still matters in classification and sorting. A recent report on the “strategic role of identity resolution” by the consulting firm Forrester suggests that the “ability to accurately identify customers” constitutes the “most basic prerequisite for marketing analytics, orchestration, and execution”. Thus, companies should combine “multiple sources of identifier and interaction information” in order to “build robust customer profiles based on multiple data sources and interactions”.[27] Accordingly, the consumer data broker Acxiom has rebranded itself as an “identity resolution” company.[28] When Roger Clarke wrote about

[22] Ibid.

[23] Ibid.

[24] See e.g. Christl (2017), p. 38-39

[25] Gandy, Oscar H. (1993): The panoptic sort: A political economy of personal information. Boulder: Westview.

[26] Lyon, David (2003): Surveillance as social sorting: Computer codes and mobile bodies. In: Lyon, D. (Ed.): Surveillance as social sorting: Privacy, risk, and digital discrimination. Routledge, London, New York.

[27] Stanhope, Joe; Mary Pilecki; Fatemeh Khatibloo; Tina Moffett; Arleen Chien; Laura Glazer (2016): The Strategic Role Of Identity Resolution. Identity Is Context In The Age Of The Customer. Forrester, October 17, 2016.

[28] In the HTML title tag of their website they use the phrase “Identity Resolution – Acxiom”: https://www.acxiom.com [26.09.2017]

dataveillance in 1988, he already emphasized the importance of “identification schemes”, as well as of the power of **data matching**, wherein different organizations cross-reference data pertaining to a large number of people “into systems that can function as the hub of a data-interchange network”.[29]

While he clearly recognized that pervasive dataveillance does not need to happen in a centralized manner, he could of course not foresee the high-frequency data matching and trading that takes place in today’s online marketing data markets. Google and Facebook dominate the field[30], but thousands of advertising technology companies, as well as myriads of website publishers, app providers, and businesses across diverse industries equally contribute to and profit from today’s personal data markets. In **programmatic advertising**, dozens of vendors integrate, combine, and auction behavioral data streams from several sources within milliseconds during a single website view. Only a few companies have all the collected profile information in one place. Often, profiles about individuals are put together only for a single interaction by combining information from multiple companies in the moment upon being triggered by certain behaviors and distributed identification technologies.[31]

(De)centralized networks of digital tracking and profiling. While Google, Facebook, and other large players managing extensive digital profiles about billions and could therefore be seen as systems of centralized mass dataveillance, decentralized networks of digital tracking and profiling that collaboratively capture every interaction across the digital world have emerged as well. In this way, companies can find and target users with certain characteristics or behaviors, learn more about them, assess them, follow them, and measure how they react, including on websites, platforms, and devices that they do not control themselves. The digital profiles they process are not static, but dynamic; they may be not comprehensive, but fragmented and distributed across several databases.[32] Due to the extensive use of inferred and predicted characteristics and behaviors these profile are often not accurate, but rather consist of estimations.[33] However, data companies constantly aim to improve data quality;[34] furthermore, many corporate databases that contain much more traditional hard facts on individuals, from banks, insurers, telecoms, and data brokers to the new platform sovereigns. These heterogeneous realms of data wealth are rapidly joining forces.[35]

[29] Ibid.

[30] http://adage.com/article/digital/verizon-chases-digital-duopoly-facebook-google/305258/

[31] Christl (2017), p. 44-46

[32] Christl (2017), p. 40-53

[33] See e.g. Schiff, Allison (2017): More Than Half Of Age Data In Mobile Exchanges Is Inaccurate. AdExchanger, January 11, 2017. Available at: https://adexchanger.com/data-exchanges/half-age-data-mobile-exchanges-inaccurate/

[34] For example, by systematically aggregating identifiers in order to link and combine profiles in more reliable ways, see e.g. Stanhope, Joe; Mary Pilecki; Fatemeh Khatibloo; Tina Moffett; Arleen Chien; Laura Glazer (2016): The Strategic Role Of Identity Resolution. Identity Is Context In The Age Of The Customer. Forrester, October 17, 2016.

[35] Christl (2017), p. 79-83

2. Personal data industries

Corporate personal data collection has long been distinguished into two fields of application: one focuses on credit and risk assessment; the other relates to direct marketing and customer data management. While the former deals with data-driven decisions that may carry very serious consequences for individuals, the latter was always much less regulated. Based on previous research[36] by the author, this chapter summarizes current corporate practices and recent developments in those two – increasingly overlapping – areas. In addition to the marketing and risk data industries, it covers the role of the new large platforms and other powerful centralized players in today's personal data industry.

2.1 Credit and risk assessment

The risk data industry consists of both largely centralized general-purpose credit reporting agencies as well as specialized companies in the fields of identity verification, employment and tenant screening, fraud prevention and detection, and insurance analytics. In countries in which these companies are allowed to cover broad areas of life, such as the US, their extensive private population data registries and automated systems have wide-reaching consequences for everyone's opportunities and life chances. Their data is supplied by banks, lenders, collection agencies, insurers, utility and telecom providers, postal services retailers, and many other kinds of institutions that capture information pertaining to essential aspects of life.[37]

In recent years, the risk data industry has massively expanded into the digital world. Online fraud detection and cybersecurity services monitor and evaluate billions of digital transactions per day. They have started to link these vast amounts of digital information to offline identity and risk assessment data. This process also blurs the boundaries between commercial risk analytics companies, law enforcement, and government surveillance. New players in financial services are testing the expansion of credit and risk monitoring to every aspect of someone's life by including behavioral data such as phone calls, browser history, social media activity, social relationships, and movement. Most major firms in business software, analytics, and consulting also play a significant role in managing and analyzing personal data for insurers, banks, and governments; examples of such companies include IBM, Informatica, SAS, FICO, Accenture, Capgemini, Deloitte, and McKinsey, and intelligence and defense firms such as Palantir.[38]

2.2 Marketing and customer data

Today's marketing and customer data industry consists of centralized general-purpose consumer data brokers and specialized companies in the fields of customer data management,

[36] Christl and Spiekermann (2016), Christl (2017)
[37] For this section see: Christl (2017), p. 27-39
[38] Ibid.

data aggregation, and analytics. The marketing and customer databases of businesses in all industries also contribute to and profit from extensive surveillance. These include sectors such as retail, consumer goods and services, travel, media and publishing, telecom and device providers, and finance. The processing of personal data for purposes of marketing and customer management is much less regulated than it is for risk management. Nevertheless, otherwise more strictly regulated sectors such as banking, insurance, and telecom extensively collect and share data for marketing purposes. In addition, these actors are lobbying heavily for the chance to exploit transactional data such as call records and payments. Charities, organizations with the goal of shaping political decisions, and parts of the public sector are likewise invested into the marketing data realm.[39]

Over the last ten years, the rise of social media, smartphones, and online advertising has embedded the collection and utilization of digital information about consumers into many areas of life. The pervasive real-time surveillance machine developed for online advertising is now rapidly integrating with long-established practices of consumer segmentation and database marketing. Today, companies can find and target users with specific characteristics and behaviors in real-time, regardless of which service or device they used, which activity they pursued, or where they are located at a given moment. Within milliseconds, these systems auction and sell digital profiles about consumers to the highest bidder. The personal information used to achieve this is not only managed by platforms such as Facebook and Google and large consumer data brokers, but also by decentralized networks of digital tracking and profiling that consist of a wide range of advertising technology, data, and analytics companies. Website publishers and app developers also provide user data on a massive scale, as do other industries that sell products and services to consumers. It is this latter group that eventually makes the most use of the digital profiles produced by such processes.[40]

Many businesses – along with other entities, such as political campaigns – can now easily utilize the data companies' services to recognize, link, and match people across different corporate databases and combine data about offline purchases with online behaviors. They can seamlessly collect rich data about consumers, add additional information, and utilize the resulting enriched digital profiles across a wide range of technology platforms.[41]

2.3 Large platforms and centralized players

Google and Facebook, followed by other large tech companies such as Apple, Microsoft, Amazon, and Alibaba, have unprecedented access to data pertaining to billions of peoples' lives. Although they have different business models and accordingly play different roles in today's personal data industry, they wield the power to dictate the basic parameters of the overall digi-

[39] For this section see Christl (2017), p. 40-53

[40] Ibid.

[41] Ibid.

tal markets in a wide-ranging manner. These large platforms enact restrictions on how other companies can obtain their data, often forcing them to utilize the wealth of the platforms' user data within the bounds of those ecosystems. Simultaneously, though, these large players also acquire additional data from beyond their platforms' reach.[42] Their role as identity providers[43] for billions and their wealth in personal data contribute to and solidify their excessive market dominance.[44] In this way, large platforms profit from economies of scale and network effects[45].

In spite of this dominance of the platforms, some of the traditional industry players are in an excellent position to join the game on a large scale – or have already done so. The large media conglomerates have deeply embedded themselves into today's tracking and profiling ecosystems; in many cases, they have even developed or acquired data and tracking capabilities themselves. For example, Time Inc. acquired a major cross-device tracking and advertising technology firm,[46] as well as a company claiming to have "access to over 1.2 billion registered users".[47] With Comcast acquiring NBC Universal, and AT&T most likely acquiring Time Warner, the large telecoms in the US are also becoming giant publishers, creating powerful portfolios of content and data capabilities. With its acquisition of AOL and Yahoo, Verizon arguably also turned into a "platform".[48] Other telecom companies such as the Norway-based Telenor[49] or the Singapore-based Singtel[50] have also acquired data technology companies already engaged in tracking billions of devices and people. The old credit card giants Visa and MasterCard have been referred to as "the first real modern platform monopolies built on big data".[51]

With regards to "artificial intelligence" approaches to analytics, which widely depend on access to vast amounts of data, Ryan Calo noted that in 2017 "as few as seven for profit institutions—Google, Facebook, IBM, Amazon, Microsoft, Apple, and Baidu in China—hold AI capabilities that vastly outstrip all other institutions".[52]

[42] Christl (2017), p.

[43] See e.g. Mirani, Leo (2014): How Facebook and Google are taking over your online identity. Quartz, September 26, 2014. Available at: https://qz.com/271286/how-facebook-and-google-are-taking-over-your-online-identity/

[44] See e.g. Me, my data and I: The future of the personal data economy. DECODE report, September 2017. Available at: https://decodeproject.eu/publications/me-my-data-and-ithe-future-personal-data-economy, p. 25

[45] See e.g. Hagiu, Andrei and Wright, Julian (2015): Multi-Sided Platforms. International Journal of Industrial Organization, Vol. 43, 2015. Available at: https://ssrn.com/abstract=2794582

[46] https://adexchanger.com/ad-exchange-news/time-inc-acquire-adelphic-build-people-based-dsp/

[47] http://www.adelphic.com/2017/01/time-inc-s-viant-acquire-adelphic/ [25.04.2017]

[48] http://www.cnbc.com/2016/10/24/how-atts-time-warner-deal-strengthens-its-position-in-advertising.html

[49] https://www.tapad.com/device-graph/ [26.04.2017]

[50] https://adexchanger.com/online-advertising/singtels-amobee-snaps-turn-310m/

[51] Stoller, Matt (2017): Equifax Isn't A Data Problem. It's A Political Problem. Huffington Post, 09/13/2017. Available at: http://www.huffingtonpost.com/entry/equifax-credit-bureaus-reform_us_59b95627e4b0edff97187e7d

[52] Calo, Ryan (2017): Artificial Intelligence Policy: A Roadmap (August 8, 2017). Available at: https://ssrn.com/abstract=3015350

2.4 Key developments in recent years

In 2007, Apple introduced the smartphone, Facebook reached 30 million users, and companies in online advertising started targeting ads to Internet users based on data about their individual preferences and interests.[53] Ten years later, a rapidly growing number of our interactions and behaviors are monitored, analyzed, and assessed by a network of machines and software algorithms operated by well-known tech giants and businesses providing products and services, as well as by myriads of companies people rarely hear of. Consumers face a situation in which thousands commercial institutions constantly record, store, and share personal information about them.[54]

One of the major developments in recent years is that companies can now address, identify, and recognize consumers on an individual level across a growing number of disparate situations in their lives. They increasingly aggregate data suitable for combining, linking, and cross-referencing profile data from different sources.[55] In particular, the pervasive real-time surveillance machine developed for online advertising is rapidly expanding into other fields, from customer data management and personalization to pricing and risk management.[56]

As a result, many remaining dams and barriers between data about “offline” behaviors, all sorts of customer data records, risk assessment information, and data recorded on the web, mobile and by many other kinds of devices have been broken. The following list summarizes some of the key areas in recent years that have contributed to this development:

- **Large platforms vs. other businesses.** Until recently, marketers who used Facebook, Google, or other online ad networks could only target individual profiles based solely on online behavior. In 2012, Facebook started to allow companies to upload and match their own lists of email addresses and phone numbers from their customer databases to its platform.[57] This lets companies systematically connect their own customer data with Facebook’s data. Moreover, it allows also other advertising technology vendors to synchronize with the platform's databases and tap into its capacities, essentially providing a kind of real-time remote control for Facebook’s data environment. Google[58] and Twitter[59] launched similar features in 2015.
- **Customer databases vs. the digital sphere.** Not only large platforms allow companies combining online and “offline” data. Since a few years, businesses in diverse industries

[53] Christl and Spiekermann (2016), p. 118
[54] See Christl (2017)
[55] Christl (2017), p. 67-69
[56] Christl (2017), p. 40-53
[57] Facebook introduced its “custom audiences”: http://techcrunch.com/2012/10/11/facebook-custom-audience-ads/
[58] Google introduced “custom match”: https://adexchanger.com/mobile/google-allows-targeted-ads-based-on-first-party-data/
[59] Twitter introduced its “partner audiences”: http://venturebeat.com/2015/03/05/twitters-new-partneraudiences-will-help-more-advertisers-track-you-outside-twitter/

can use “data management platforms” and other technology services as central hubs that aggregate, integrate, manage, and deploy different sources of data about consumers, including but not limited to data that companies have been collecting themselves.[60] Companies have begun linking data from the web and smartphones with the customer data and offline information that they have been amassing for decades.

- **Payment data as a major bridge.** In recent years, credit card networks have started to make data about their customers’ purchases available to the digital tracking and profiling universe. Google stated that it captures “approximately 70% of credit and debit card transactions in the United States” through “third-party partnerships” in order to track purchases.[61] But information about credit card interactions is also available to other companies, for example via consumer data brokers such as Oracle.[62] Analysts have stated that for MasterCard, selling products and services created from data analytics might even become its “core business” given that “information products, including sales of data” already represent a considerable and growing share of its revenue.[63]
- **Risk assessment data and marketing.** Key players in data, analytics, and technology that provide risk assessments of individuals in important fields of life such as credit and insurance mostly also provide marketing solutions. Other companies use their data to sort, rank, target, or exclude consumers based on their estimated profitability.[64] As perhaps the most extreme example of this, Twitter ads can now be targeted by “creditworthiness”, thanks to data provided by Oracle.[65]
- **Risk assessment based on everyday life behaviors.** Smaller “fintech” companies have started to predict consumers’ creditworthiness based on factors such as the timing and frequency of phone call records, GPS location, customer support data, online purchases, web searches, and data from social networks, including information about someone’s social network connections.[66] Along similar lines, nobody knows whether Facebook will turn its patent for assessing creditworthiness based on someone’s friends[67] into reality.
- **Insurance programs incorporating data about everyday life behaviors.** Large insurers in the US and in Europe have introduced programs that allow consumers to get significant discounts on their insurance premiums if they agree to provide real-time data about car driving behavior and activities such as their steps, grocery purchases, and fitness studio visits.[68]

[60] Christl (2017), p. 48
[61] https://adwords.googleblog.com/2017/05/powering-ads-and-analytics-innovations.html
[62] Christl (2017), p. 60
[63] http://paymentweek.com/2014-6-16-for-mastercard-processing-and-analytics-go-hand-in-hand-4908/
[64] Christl (2017), p. 79-83
[65] https://twitter.com/WolfieChristl/status/850467843430912000
[66] Christl (2017), p. 30-31
[67] http://www.theatlantic.com/technology/archive/2015/09/facebooks-new-patent-anddigital-redlining/407287
[68] Christl and Spiekermann (2016), p. 52-68

- **Calculating health risks based on consumer data.** In healthcare, data companies and insurers are working on programs that use everyday life data about consumers to predict someone's health risks.[69] Data and analytics companies have started to offer health scoring products that predict individual health risks of people based on vast amounts of consumer data, including purchase activities.[70]
- **Online fraud detection is connecting the dots.** The ubiquitous streams of behavioral data in the digital world are also being fed into fraud detection systems, which use highly invasive technologies to evaluate billions of digital transactions and collect vast amounts of information about devices, individuals, and suspicious behaviors. Companies have started combining information about devices, online behaviors, and digital transactions with personal identity and credit information.[71] ID Analytics, a US-based credit and fraud risk data company recently acquired by Symantec, runs an "ID Network" with "100 million identity elements coming in each day from leading cross-industry organizations"[72], containing data about 300 million consumers.[73] TransUnion's Trustev even offers "social fingerprinting", which includes "friend list analysis" and "pattern identification" analyzing social media content. The latter, at least, is only used with individuals' "full permission" through a "voluntary social network login".[74]

These examples show how information about people's behaviors, social relationships, and most private moments is increasingly applied in contexts or for purposes completely different from those for which it was recorded. Most of these developments have happened during the last few years and constitute the background for the further examination of commercial data uses.

[69] Christl (2017), p. 80

[70] http://www.lexisnexis.com/risk/downloads/literature/health-care/socioeconomic-data-coverages-br.pdf [02.05.2017]

[71] Christl (2017), p. 34-38

[72] http://www.idanalytics.com/data-and-technology/idnetwork/ [23.04.2017]

[73] http://www.idanalytics.com/media/VA-Resolve360-Datasheet.pdf [23.04.2017]

[74] Trustev Sales Pack. REAL TIME ONLINE IDENTITY VERIFICATION. PDF Brochure. Personal copy of file with author Wolfie Christl

3. When personal data is systematically used against people

While data-driven technologies can produce immense benefits for everyone in many areas of life, companies and other institutions can easily use their data wealth against people. The possible adverse effects of corporate data collection and utilization on individuals, groups of people, and society are diverse, but rarely considered in the commercial sphere. This chapter explores and examines corporate practices and its societal implications in two – in part overlapping – areas of concern, automated decision-making and data-driven persuasion.

3.1 Automated decisions based on personal data

As Frank Pasquale and Danielle Citron have stated, data is increasingly used to "assess whether we are good credit risks, desirable employees, reliable tenants, valuable customers — or deadbeats, shirkers, menaces, and 'wastes of time'".[75]

Automated decisions based on personal information and analytics can have serious consequences for people; they may affect people's choices and life-chances, and, on a fundamental level, their general autonomy and human dignity. The resulting effects may be distributed unevenly across different population groups and accumulate over time. This section focuses mostly on how personal information is used in situations where powerful commercial entities make automated decisions about people, individuals, consumers, and citizens who are in less powerful positions.[76] Those decisions may happen either directly or indirectly, for example, when banks and insurers utilize credit scores or similar data provided by other powerful commercial parties.

Example areas and types of application

At their core, automated decision-making systems exist in order to treat people differently on the basis of information about them. As a result, individuals get excluded from certain opportunities, become subject to further investigation, or are filtered out in advance.

Eligibility decisions for a loan, other financial services, insurance, healthcare, housing, education, or employment can have significant and immediate impacts by excluding people outright. Equally significant economic effects can stem from less favorable service, terms, or prices, such as through fees, interest rates, or insurance premiums.[77] Data-driven decisions may

[75] Citron, Danielle Keats and Frank A. Pasquale (2014): The Scored Society: Due Process for Automated Predictions. Washington Law Review, Vol. 89, 2014, U of Maryland Legal Studies Research Paper No. 2014-8, p. 1. Available at: https://ssrn.com/abstract=2376209

[76] For an overview of the relationship between powerful parties and data subjects see e.g. Rhoen, M. (2016): Beyond consent: improving data protection through consumer protection law. Internet Policy Review, 5(1). Available at: https://policyreview.info/articles/analysis/beyond-consent-improving-data-protection-through-consumer-protection-law

[77] Valentino-DeVries, Jennifer; Jeremy Singer-Vine snd Ashkan Soltani (2012): Websites Vary Prices, Deals Based on Users. Wall Street Journal, Dec. 23, 2012. Available at: http://on.wsj.com/Tj1W2V; Wells Fargo, for example, it steered an estimated 30,000 black and Hispanic borrowers from 2004-2009 into more costly subprime loans or charged them higher fees than

either occur in a **fully automated** manner, as in the case of a bank account or credit application denial, or they may happen prior to the actual decision, for example, when unwanted people are automatically rated low and **filtered out**, and thus never seen by human staff or by a system farther on in the process. Short of being denied outright, people may become **subject to further investigation** because they received a low score that might, for instance, have suggested that a person has a higher likelihood of committing fraud.[78] As a result, even before an eligibility decision is conducted, an applicant may need to disclose further personal information, undergo complicated bureaucratic processes, or, for instance, submit further medical exams in order to avoid becoming excluded.[79] In the process, applicants may drop out deterred or unable to comply, simply because they are facing too many competing demands, or because more errors have crept into the process that ultimately doom the applicant. Similarly, even after an application was approved for a loan successfully, a bank might automatically flag a customer when its system somehow predicts, based on behavioral data, a high risk of default many months in advance, subjecting the individual to extra scrutiny and higher costs as a result.[80]

Grey areas. Because of its potential to significantly shape people's lives, the corporate use of personal data in most of the areas mentioned above has been regulated in many regions of the world.[81] But many of today's uses take place either in grey areas or in wholly unregulated territory. Financial services companies, for example, use "prescreening" or "prequalification" systems based on credit data and consumer behavior to select customers and deliver pitches during a personal consultation or via email or phone, and then only present a certain selection of products and services to individuals.[82] Identity verification and fraud prevention systems do not only decide whether someone is suspicious or not, but may rank-order accounts and prioritize less risky customers to optimize costs.[83] Moreover, both marketing efforts and how customers are treated by a company are directed towards certain groups of people and therefore exclude others, based on how relevant these groups are rated by the company's customer relationship management systems.[84]

comparable white borrowers. O'Toole, James (2012): Wells Fargo in $175M discriminatory lending settlement. CNN Money, July12, 2012. Available at: http://money.cnn.com/2012/07/12/real_estate/wells-fargo-lending-settlement/

[78] See e.g. Christl (2017), p. 32-34

[79] Until the first half of the 20th century, US life insurers generally sold substandard plans to minorities and required them to submit additional medical exams, see e.g. Angwin et al (2017): Minority Neighborhoods Pay Higher Car Insurance Premiums Than White Areas With the Same Risk. ProPublica, April 5, 2017. Available at: https://www.propublica.org/article/minority-neighborhoods-higher-car-insurance-premiums-white-areas-same-risk

[80] Kennedy, K., Mac Namee, B., Delany, S. J., O'Sullivan, M., & Watson, N. (2013): A window of opportunity: Assessing beha v-ioural scoring. Expert Systems with Applications: An International Journal, 40(4), 1372-1380. Available at: http://arrow.dit.ie/cgi/viewcontent.cgi?article=1024&context=scschcomart

[81] For some examples see: Christl (2017), p. 27-30

[82] Christl (2017): Corporate Surveillance in Everyday Life, p. 81

[83] Ibid.

[84] Christl (2017), p. 40-57

The more digital technology and personal data collection become part of everyday life, the more pervasive and opaque such practices become. While companies have used consumer's zip codes to decide whether or not to market certain products or services to them in the past,[85] today they may use digital records about many other kinds of attributes and behaviors of consumers to make those decisions. Apart from discriminating against people by providing some with more expensive offers than others – or by excluding them outright – there are many further ways to underserve people and keep them away, ranging from pre-filtering marketing and advertising to prioritization in call centers or ticketing systems.[86] Conversely, companies may focus on customers with a tendency to incur late payment costs or other penalties.[87]

The possibilities for practically implementing such business objectives are growing exponentially, with much of today's online marketing driving the process. Increasingly consumer decisions are shaped by a sophisticated automated one-to-one direct marketing process[88] that includes real-time behavioral targeting, retargeting, personalized offers, and customized discounts down to the level of specific individuals. For this purpose, companies can not only personalize their own environments and channels such as their website, but also monitor consumers across myriads of other websites, platforms, services, and apps throughout the digital world.[89]

Many smaller disadvantages. Single automated decisions based on personal data can have far-reaching consequences, such as when a bank, insurer, telecom, or energy provider simply denies service to someone. In other cases, being rated as an unwanted, suspicious, or non-valuable person may lead to being systematically excluded or disadvantaged on a smaller scale, but many times. The latter may include someone being automatically denied registration for a service or ordering at an online shop, but also, for instance, not being offered certain payment and shipping options or receiving worse conditions regarding returns. Based on constant digital tracking and profiling certain individuals may wait longer when calling the customer service line than others, get different personalized offers and discounts, see differently priced products, or even get different prices for the same products.[90]

Not getting the option to pay with a credit card in an online shop a single time certainly does not carry the same implications for an individual as being denied a bank account or job. Simi-

[85] See e.g. Angwin et al (2017): Minority Neighborhoods Pay Higher Car Insurance Premiums Than White Areas With the Same Risk. ProPublica, April 5, 2017. Available at: https://www.propublica.org/article/minority-neighborhoods-higher-car-insurance-premiums-white-areas-same-risk

[86] See e.g. Cathy No'Neil (2016): Weapons of Math Destruction, p. 143

[87] Oscar H. Gandy, Jr. (2012): Statistical surveillance: Remote sensing in the digital age. In K. Ball, K. Haggerty and D. Lyon (eds), Routledge Handbook of Surveillance Studies. New York, Routledge, 2012, p. 127

[88] In contrast to classic advertising that is addressed to large groups of people, online advertising is increasingly targeted and addresses to specific individuals based on their previous behaviors, similar to direct marketing of the past. See Christl (2017) or e.g. http://zgp.org/targeted-advertising-considered-harmful/

[89] Christl (2017), p. 40-57

[90] Christl and Spiekermann (2016), p. 125

larly, not seeing a single ad may of course not have any negative consequences for an individual; many might even be happy to be left alone. But when these occurrences happen in a systematic way, cumulative disadvantage[91] can lead to differential access to information, opportunities, and chances across society.

Deciding about people based on their data

A system making automated decisions takes different personal attributes related to a person as input data; these may include someone's age, zip code, income, credit history, medical history, or, perhaps someone's browser history. Such a system may be based on anything from very simple rules to complex statistical models performing different tasks such as classification, estimation or prediction.[92] In the case of classification, each individual is, depending on the input data, put into one of several classification[93] categories; a system may, for example, judge someone as a "good" or "bad" borrower. In the case of prediction[94], the system assigns each affected individual a number that expresses the likelihood of certain future actions or states, for example, the future ability to repay a loan or the future health status.

To develop or continuously adapt these decision models, a system typically also uses different kinds of non-personal information and personal data from others.[95] From a privacy perspective focusing on the individuals affected by an automated decision, however, crucial questions include: **which kinds of personal information** are used as input data in the moment of application? Was said data deliberately volunteered by the person, observed without the person's full knowledge, or perhaps acquired from a third-party, whether a private one or a publicly available source?[96]

Moreover, a data-driven decision can happen with or without the fully informed **consent** and **knowledge** of the subject. Individuals may or may not have a reasonable level of **choice** to be subject of the automated decision at all. Furthermore, it may not be possible to **object** to such a decision. The reasons for a resulting decision may or may not be **explainable** to an individual. The system's objectives, functionalities, accuracy and impact on groups and society may or may not be **transparent** or subject to external inspection, evaluation, or auditing. The decision-making company or institution may or may not be being held **accountable and liable** for such a decision.

[91] See e.g. Gandy, Jr, Oscar. (2009): Coming to terms with chance: Engaging rational discrimination and cumulative disadvantage. Coming to Terms with Chance: Engaging Rational Discrimination and Cumulative Disadvantage. 1-240.
[92] See e.g. Linoff, Gordon S. and Michael J. A. Berry (2004): Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management. Wiley Publishing.
[93] Ibid.
[94] Ibid.
[95] Ibid.
[96] Christl (2017), p. 15

Fields of application such as the use of traditional credit scores for automated decisions in financial services, which have long been regulated in some way, show massive problems regarding accuracy[97], arbitrariness, opacity, and the disparate impacts they have on different groups of people.[98] The application of data-driven decisions in today's digital environments – from disabled user accounts[99] to personalized pricing[100] to denied payment or shipment methods in online shops – is barely examined. Existing research – on, for example, underlying systems such as online fraud detection and scoring[101] – suggests that the resulting decisions are mostly nontransparent and opaque, occur without the fully informed consent and knowledge of users, use data from a wide range of sources in an uncontrolled way, are not explainable to users and cannot be objected by them, are not externally evaluated or audited, and are not sufficiently (if at all) subject to accountability mechanisms for inaccurate decisions. Many other fields of application have already – or may soon – become relevant in evaluating the digital footprints of individuals' everyday lives, from financial service providers[102] to digital work platforms[103].

From the perspective of an individual that is subject to an automated decision, several reasons can lead to being treated differently than others based on his or her personal information:

- **Technically accurate decisions.** Someone may be treated differently than others based on personal data because a system is working exactly as intended. When a system designed to deny people with a yearly income below $30,000 accurately ascertains that this is the case for someone on the basis of the input data, this is obviously an accurate decision, at least on a technical level.
- **Inaccurate input data.** For a variety of reasons, much of the data collected on people is inaccurate. Systemic or procedural flaws may occur in its collection, aggregation, matching, and transfer. Data may originate from unreliable sources, or may be incomplete or outdated. Software contains bugs.[104] Even credit reports, which are the basis for credit scoring and amongst the best regulated realms in commercial data collection, often contain serious errors.[105]

[97] Christl (2017), p. 29

[98] Citron, Danielle Keats and Frank A. Pasquale (2014): The Scored Society: Due Process for Automated Predictions. Washington Law Review, Vol. 89, 2014, U of Maryland Legal Studies Research Paper No. 2014-8, p. 1. Available at: https://ssrn.com/abstract=2376209

[99] User accounts at platforms such as Facebook, Google or Amazon increasingly become important for people's economic lives, see e.g. https://www.theguardian.com/money/2016/mar/18/banned-by-amazon-returning-faulty-goods-blocked-credit-balance

[100] Christl and Spiekermann (2016), p. 41-44

[101] Christl and Spiekermann (2016), p. 38-40; Christl (2017), p. 34-38

[102] See section 2.1

[103] See e.g. Calo, Ryan and Rosenblat, Alex (2017): The Taking Economy: Uber, Information, and Power (March 9, 2017). Columbia Law Review, Vol. 117, 2017; University of Washington School of Law Research Paper No. 2017-08. Available at: https://ssrn.com/abstract=2929643

[104] Christl and Spiekermann (2016), p. 126

[105] See Christl (2017), p. 29 and Peppet, Scott R. (2019): Unraveling Privacy: The Personal Prospectus & the Threat of a Full Disclosure Future (Aug 7, 2010). Northwestern Univ. Law Review, 2011, p.1178. Available at: https://ssrn.com/abstract=1678634

- **Inaccurate decisions.** Classifying or ranking people – not to mention predicting their future behavior – based on personal information can go wrong in many ways. Analysis methods based on correlations and probabilities are far from objective; rather, their blurriness is inherent to their design.[106] As a result, someone who knows the wrong people, lives in the wrong district, visits the wrong shop, or surfs the wrong website may get categorized and judged in a certain negative way. Companies may de-contextualize and misinterpret recorded consumer interactions. In general, the underlying "motivations for particular actions are never explained or understood" by these systems.[107]
- **Technically accurate, non-intentionally biased decisions.** In this case, inferences are not inaccurate; rather, they are biased against certain groups of people – such as those of certain genders or ethnicities – but without this being intended by a company's business objectives. These biases and flaws in the data mining process may be caused by e.g. wrongly choosing variables that correlate to certain groups more than they do to others. Generally, these systems may reproduce prejudices of prior decision-makers, unaware engineers, or preexisting patterns of exclusion and inequality in society at large.[108]
- **Technically accurate, intentionally biased decisions.** Companies may intentionally discriminate against certain groups of people by directly including information about someone's gender, age, ethnicity, religion, political opinion, health status, sexual orientation, or socioeconomic status into these decisions. Alternatively, they may infer someone's membership in one of these groups through proxy attributes that correlate with group membership. This may be difficult to detect.[109] Data-driven decisions could "breathe new life into traditional forms of intentional discrimination",[110] not in the least because they can access attributes referring or correlating to group membership much more efficient and on a large-scale.
- **Too little or no data.** Refusing to share personal information or participate in today's digital tracking may have consequences, too. If not enough data about a person is available, the risk of a customer relationship may be considered as too high up-front. In this way, automated decisions can privilege those groups of people more willing – or able – to participate in the data-driven aspects of contemporary society.

In sum, when someone is denied by an automated system on the basis of his or her data this may occur because of a technically accurate decision, but also because of inaccurate data, too

[106] See e.g. danah boyd & Kate Crawford (2012): CRITICAL QUESTIONS FOR BIG DATA, Information, Communication & Society, 15:5, 662-679. Available at: http://dx.doi.org/10.1080/1369118X.2012.678878

[107] De Zwart, Melissa; Humphreys, Sal; Van Dissel, Beatrix (2014): Surveillance, big data and democracy: lessons for Australia from the US and UK, UNSW Law Journal. Available at: http://www.unswlawjournal.unsw.edu.au/sites/default/files/final_t3_de_zwart_humphreys_and_van_dissel.pdf

[108] Solon Barocas and Andrew Selbst (2016): Big Data's Disparate Impact, California Law Review, Vol. 104, 2016. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899

[109] Barocas, Solon (2014): Data Mining and the Discourse on Discrimination. Conference on Knowledge Discovery and Data Mining. Available at: https://dataethics.github.io/proceedings/DataMiningandtheDiscourseOnDiscrimination.pdf

[110] Solon Barocas and Andrew Selbst (2016)

little or no data, or because of generally flawed or otherwise – whether intentionally or non-intentionally – biased decision-making. Automated decisions based on personal data can definitely have serious consequences for individuals who are subjects to such decisions. When they are systematically biased against certain groups, this has implications for society at large.

Some cases are easier to address than others. Where discriminatory decisions against certain groups, whether intentional or not, are forbidden due to anti-discrimination law or data protection legislation, such must be enforced. In order to do so, existing systems must urgently be made much more transparent. Because companies rarely increase transparency of their own volition, additional legislation will be required. Moreover, research about non-intentional biases must be intensified. Policy efforts and public debate on how to make existing and future data-driven decision-making systems fair, accountable, and transparent (FAT) are needed as well.[111]

However, the trend of companies increasingly turning individuals into "ranked and rated objects"[112] leads to further problems on a more fundamental level.

Self-fulfilling prophecies. Automated decisions may not only reproduce, but also reinforce and worsen inequality. Credit scores provide a good example of this. They may create the "financial distress they claim merely to indicate", thus becoming self-fulfilling prophecies.[113] The "act of designating someone as a likely credit risk" may raise the cost of future loans or insurance rates or decrease said individual's employability.[114] The consumer-finance scholar Federico Ferretti fundamentally questions the "capability of credit data to prevent future over-indebtedness", because one cannot foresee many major causes of over-indebtedness, such as illnesses, divorce, job losses, and poor market conditions through it.[115]

Kafkaesque experiences and chilling effects

More than the actual decision itself, the mere fact of being subject to an automated decision based on personal data may already have significant effects on people. The awareness – whether real or imagined – that data about behavior and personality will be used to determine one's future environments, opportunities, and chances, may make someone feel and act differently.

[111] See e.g. conferences such as "Conference on Fairness, Accountability, and Transparency" (https://fatconference.org) or "Fairness, Accountability, and Transparency in Machine Learning" (http://www.fatml.org)

[112] Citron, Danielle Keats and Frank A. Pasquale (2014): The Scored Society: Due Process for Automated Predictions. Washington Law Review, Vol. 89, 2014, U of Maryland Legal Studies Research Paper No. 2014-8, p. 3. Available at: https://ssrn.com/abstract=2376209

[113] Ibid, p. 18

[114] Ibid.

[115] Ferretti, Federico (2015): Credit Bureaus Between Risk-Management, Creditworthiness Assessment and Prudential Supervision. EUI Department of Law Research Paper No. 2015/20. Available at: https://ssrn.com/abstract=2610142

One can distinguish between two "ideal" types of data-driven decisions:

1. Opaque ones that may even occur invisibly, i.e. without the subject's knowledge.
2. Transparent ones that occur with the fully informed consent of the subject of the decision, who also knows how the decision is influenced by the provided personal information.

The first case is the norm in today's digital world. Whether online fraud detection, user account validation, behavioral advertising or pricing, automated decisions in these areas are typically opaque and occur invisibly. In most cases, only the companies providing these systems know which data they are based on.[116] There are rarely ways to object to decisions, and if there are, processes are bureaucratic and unlikely to lead to satisfying results.[117] In the worst case, the affected person might not even know that he or she has been subject to a negative decision, because better offers or certain options are simply not available. This may lead to Kafkaesque experiences,[118] particularly in light of all the contemporary practices of data-driven personalization and customer management, which this paper examines further in section 3.2. Today, every click on a website and every swipe on a smartphone may trigger a wide variety of hidden data sharing mechanisms distributed across several companies and, as a result, directly affect a person's available choices. As a result, consumers never know whether their everyday life behaviors may lead to a response from any of those continuously updated, interconnected, opaque data networks, and, if so, how this affects the content they see and the options they are given across many life situations.

The second case – fully transparent and comprehensible automated decisions – is virtually non-existent. Even credit decisions are opaque and arbitrary.[119] Although consumer reporting agencies, as the director of the US Consumer Financial Protection Bureau recently stated, exert a "tremendous influence over the ways and means of people's financial lives"[120], credit scores are nontransparent, the algorithms used to calculate them are secret, and, at least in the US, there are no external audits.[121] Although the categories' relative weights are known, consumers remain in the dark as to how exactly their individual behaviors affect their credit scores.[122]

[116] Generally, see: Christl and Spiekermann (2016)

[117] E.g. https://www.theguardian.com/money/2016/mar/18/banned-by-amazon-returning-faulty-goods-blocked-credit-balance

[118] As described by Daniel Solove and others, see e.g. Solove, Daniel J. (2004): The Digital Person: Technology and Privacy in the Information Age (October 1, 2004). Daniel J. Solove, The Digital Person: Technology and Privacy in the Information Age, NYU Press (2004); GWU Law School Public Law Research Paper 2017-5; GWU Legal Studies Research Paper 2017-5. Available: https://ssrn.com/abstract=2899131

[119] Citron, Danielle Keats and Frank A. Pasquale (2014): The Scored Society: Due Process for Automated Predictions. Washington Law Review, Vol. 89, 2014, U of Maryland Legal Studies Research Paper No. 2014-8, p. 10-13. Available at: https://ssrn.com/abstract=2376209

[120] https://www.consumerfinance.gov/about-us/newsroom/prepared-remarks-cfpb-director-richard-cordray-consumer-advisory-board-meeting-march-2017/

[121] Ibid., p. 10-11

[122] Ibid.

Credit scores are arbitrary on an individual level, because they often contain inaccurate data, and on an aggregate level, because scores calculated by different companies provable differ.[123] Generally, consumer reporting agencies are, according to privacy schlor Chris Hoofnagle, “notoriously unresponsive and unaccountable bureaucracies”.[124] The perplexing arbitrariness of this situation has given rise to a market of books, articles, and websites offering advice on how to improve one’s credit score.[125]

Whether decisions are fully opaque or partially transparent, a consumer may have at least “a vague sense that information is being collected and used to her disadvantage, but never truly knows how or when”.[126] This can produce **chilling effects** on forms of action or expression. People may avoid behaviors that they suspect to be factors in an automated system’s judgment process. As today’s extensive digital records about everyday life behaviors increasingly determine which options and prices are available, these chilling effects become considerable. For instance, consumers might only visit a travel website once because they assume that visiting it too frequently will increase the price. Depending on what they believe to know about digital tracking and analytics they may refrain from activities such as visiting certain websites, interacting with certain contents, or searching after certain terms. They may avoid expressing themselves online or participating in public debate and generally refrain from doing anything they consider to be unwanted or nonconforming. Such effects become especially acute when extensive information about everyday life behaviors determines access to financial services, employment, and to other vital opportunities.[127]

As a result, these practices potentially limit the agency, autonomy, and dignity of individuals; this further affects society on an aggregate level. When certain groups are disproportionally more often subject of automated decisions based on their personal information, social equality suffers.

Individuals necessarily experience opaque data-driven decisions as inherently arbitrary and Kafkaesque; this leads to its chilling effects. In the as-yet hypothetical case of **fully transparent decisions**, in which individuals know exactly how information about their lives and behaviors influences the outcomes and how this information is collected or recorded, chilling effects turn into effects of digital social control. As long as the outcomes of such decisions are reasonably important, individuals will try to act correspondingly and avoid any situation or behavior that has a negative effect on the decision. As such, the specific parameters that influence the decision-making system force individuals to adapt their lives and behaviors to the require-

[123] Ibid., p. 11-12; and Christl (2017), p. 29

[124] Hoofnagle, Chris Jay (2013): How the Fair Credit Reporting Act Regulates Big Data (September 10, 2013). Future of Privacy Forum Workshop on Big Data and Privacy: Making Ends Meet, 2013. Available: https://ssrn.com/abstract=2432955

[125] Ibid, p. 11; also, a Google search for the phrase “’credit score’ explained’ returns 39 million results

[126] Calo, Ryan (2013): Digital Market Manipulation (August 15, 2013). 82 George Washington Law Review 995 (2014); University of Washington School of Law Research Paper No. 2013-27, p.1029. Available at: https://ssrn.com/abstract=2309703

[127] See e.g. Christl (2017), p. 30-31

ments of the system. This is in part also the case for partially transparent decisions, such as when the factors that influence credit scoring are known on a rough level.

As long as individuals lack a **real choice** in whether to be subject to such a – fully or in part – transparent decision, their autonomy is similarly limited – not because of uncertainty and arbitrariness, but because of direct behavioral control. In fact, individuals do not have a real choice in many cases, but are pressured into being subject to automated decisions or to disclose data, because refusing to agree would lead to serious economic or social disadvantages.

Data about everyday life. The degree to which data-driven decisions limit someone's autonomy depends on how much they may determine his or her choices and opportunities, but also on the extent of information about personality and behaviors used, as well as on how those are collected or recorded. A system that discriminates on the basis of data directly related to its goal – in the case of a credit decision, data such as amounts owed, missed payments, and bankruptcy records – will certainly affect individual autonomy less than a system that draws on extraneous data such as browser and movement histories, social media interactions, and friends lists. The latter is unlikely to be fully transparent and explainable anyway and will therefore arguably lead to uncertainty and arbitrary decisions.[128] In any case, such far-reaching use of data about everyday life introduces a level of social control that massively affects autonomy and human dignity, especially when determining vital areas of life. A system that monitors the websites visited, apps used, terms searched for on the web, places visited, or friends added on a social network in order to make discrete or cumulative decisions about a person that affect essential areas of life, could perhaps be considered as a kind of virtual imprisonment, similar to an ankle monitor that defendants under parole are required to wear.

As a way to describe the difference between personal data more or less related to automated decisions Helen Nissenbaum has introduced the concept of privacy as "contextual integrity".[129] Personal data is collected in a certain context for a certain purpose by a certain type of entity. The more the context and purpose of data use differ from the circumstances of its collection, the more the potential for negative effects on the data subjects increases.

Accurately discriminating for profit

Making decisions about individuals based on information about them – whether automated or not, whether carried out by a business or other kind of institution – constitutes a discriminatory process per se. Corporate decision makers employ such processes in order to maximize effi-

[128] Systems that incorporate such a wide range of data about everyday life are based on machine learning or other technologies that are largely opaque. Although there might be ways to improve human oversight of such systems, their classifications or predictions can be hardly explained to the subjects of a decision. See e.g. Pasquale, Frank (2016): Bittersweet Mysteries of Machine Learning (A Provocation). Available at: http://blogs.lse.ac.uk/mediapolicyproject/2016/02/05/bittersweet-mysteries-of-machine-learning-a-provocation/

[129] Nissenbaum, Helen (2004). Privacy As Contextual Integrity. Washington Law Review. 79.

ciency and increase profits. The subjects are rated, ranked, sorted into groups, and, as a result, treated differently according to the company's economic goals and business logics. This gives rise to another massive source of bias that exists on a fundamental level within corporate data-driven systems and is often overseen in the discussion of automated decisions.

Thus, rather than just examining and improving existing systems, we, as a society, should also ask: in which areas do we want private or public data-driven systems to make decisions about people, and based on which objectives and values? Who designs and controls those systems? Which kinds of behaviors should they be allowed to reward or punish? Which input data related to the personal characteristics, behaviors and lives of someone should banks, insurers, healthcare providers, employers, and other parties be allowed to use in such systems?

In banking and insurance the types of data that may be used for risk assessments are restricted by laws that reflect public policy goals, but also by the interests of industry lobby groups.[130] For example, European private life insurers use attributes such as age, occupation, and level of education, as well as behavioral habits such as smoking and drinking, to sort and group applicants according to their projected risk.[131] Similarly, motor insurers use attributes such as age, location, occupation, and claims history.[132] Banks and lenders likewise use personal information to intentionally discriminate between groups of people. Credit scoring, as defined by the consumer-finance scholar Federico Ferretti, is essentially "a way of recognising different groups in a population according to certain features, expressed by a combination of personal data and other non-personal information, and differentiating them on grounds of parameters and classifications set a priori from statistics for a predictive purpose".[133] Banks and lenders either directly use personal information such as age or income level to approve a loan[134], or they may do so indirectly by, for example, estimating age from the length of a credit history – as the latter often correlates with the former.[135]

While credit scoring was, according to Ferretti, originally intended to "minimise the percentage of consumers who default", lenders now use it to "identify the customers who are most profitable and to maximise profits through risk based pricing", while "blurring this all with direct marketing activities".[136] Credit data about borrowers is not only used "as a tool to meet the problem of asymmetrical information between borrowers and lenders" in the financial

[130] Groupe Consultatif Actuariel Européen (2011): Use of age & disability as rating factors in insurance. Why are they used and what would be the implications of restricting their use? Position Paper, December 2011. Available at: http://actuary.eu/documents/GC_Age_Disability_Underwriting_Paper_051211.pdf

[131] Ibid.

[132] Ibid.

[133] Ferretti, F. (2009): The Credit Scoring Pandemic and the European Vaccine: Making Sense of EU Data Protection Legislation, 2009(1) Journal of Information, Law & Technology (JILT). Available at: http://go.warwick.ac.uk/jilt/2009_1/ferretti

[134] Ibid.

[135] See e.g. http://www.myfico.com/credit-education/whats-in-your-credit-score/

[136] Ferretti, Federico (2015): Credit Bureaus Between Risk-Management, Creditworthiness Assessment and Prudential Supervision. EUI Department of Law Research Paper No. 2015/20. Available at: https://ssrn.com/abstract=2610142

system[137], but, reversing the information asymmetry, sensitive risk assessment data is used to maximize profits at all stages, from selective marketing, pre-screening, pricing, and terms to account management and debt collection. Ferretti argues that credit scoring should emphasize the minimization of business risk and increased profitability, which would certainly represent a legitimate business interest, but should not disproportionally limit the rights of individuals. He argues that in cases where the latter occurs, the "underlying business interests that they enhance should be limited".[138]

The increasing availability of personal data through today's pervasive digital profiling technologies has massively improved the ways companies can exploit data to treat individuals and groups of people differently for their economic advantage. While companies in financial services and insurance are working on a massive expansion of the types of personal information to be used for risk assessment and pricing – from social media and location data to physical activity recorded by wearables[139] – businesses in all industries are increasingly adopting the actuarial logic of insurers and banks. Today, companies constantly monitor, evaluate, sort, and group people in terms of how "valuable" or "risky" they might be as customers, and then treat them accordingly.[140]

Personalized discrimination. Based on the extensive amounts of personal data both collected by companies themselves and acquired or accessed from third parties, they aim to find, attract, and target valuable new customers, retain existing customers according to their profitability or "lifetime value"[141], and avoid consumers that have been classified as risky or as "waste".[142] Subsequently, consumers become subject to differential treatment, from customized telemarketing scripts, promotional materials, online content, ads, offers, discounts to pricing.[143] When, as suggested by a major consumer data broker, the top 30% of a company's customers are classified as individuals who could add 500% value, and the bottom 20% of customers are categorized as individuals who could actually cost 400% value, a company may "shower their top customers with attention, while ignoring the latter 20%, who may spend 'too much' time on customer service calls, cost companies in returns or coupons, or otherwise cost more than

[137] Ibid, p. 6

[138] Ferretti, Federico (2009)

[139] Christl (2017), p. 30; Christl and Spiekermann (2016),

[140] Christl (2017), p. 40-54

[141] In marketing, the "customer lifetime value" is a prediction of the net profit attributed to the entire future relationship with a customer. See e.g. Abdolvand, Neda; Amir Albadvi; Hamidreza Koosha (2014): Customer Lifetime Value: Literature Scoping Map, and an Agenda for Future Research. International Journal of Management Perspective, Vol. 1, No.3, pp. 41-59.

[142] Natasha Singer (2012): Mapping, and Sharing, the Consumer Genome. New York Times, June 16, 2012. Available at: http://www.nytimes.com/2012/06/17/technology/acxiom-the-quiet-giant-of-consumer-database-marketing.html

[143] Christl and Spiekermann (2016), p. 41-44

they provide".[144] Businesses may also, for example, calculate the exact minimum action necessary to keep customers loyal.

When certain consumers are constantly classified as non-valuable and risky, they may experience many small disadvantages in their everyday lives, each of them not very significant on their own, but accumulated resulting in a significant cumulative disadvantage in life. These disadvantages may result from being treated badly or excluded, but also from being selectively targeted for certain types of messages or information.

3.2 Data-driven persuasion and personalized disadvantage

Pervasive digital tracking and profiling, in combination with personalization, advanced customer management technologies, and testing, have become a powerful toolset for systematically influencing behavior.[145] Companies and other institutions can utilize vast information asymmetries to exploit personal weaknesses and cognitive biases with calculated efficiency and unprecedented effectiveness.[146] They make strategic use of these capacities in areas such as behavioral advertising, marketing and sales, as well as in news, entertainment, and political campaigning.

Some of these practices overlap with automated decision-making and the concerns about its impacts and societal implications. The distinct characteristic of the issues examined in this section is that they aim to selectively stimulate behavioral change by customizing interactions and environments, rather than to make single consequential decisions such as eligibility assessments. Data-driven persuasion capitalizes on knowledge generated from large data sets, insights from behavioral science, and from permanent experimentation on real people. Most importantly, it bases on the access to personal information at the moment of application. Data-driven persuasion can be seen as both a sub set and expansion of the practices examined in the context of automated decision-making.

User experience and behavioral advertising. Companies use persuasive technologies within closed environments such as their websites, platforms, online shops, services, apps, and software products, but also across the wider digital world. Data-driven persuasion may have

[144] Marwick, Alice E. (2013): Big Data, Data-Mining, and the Social Web. Talk for the New York Review of Books Event: Privacy, Power & the Internet, October 30, 2013, p. 5. Available at: http://www.tiara.org/blog/wpcontent/uploads/2013/10/marwick_2013_datamining_talk.pdf

[145] Christl (2017), p. 75-78

[146] Gandy, Jr, Oscar (2017): Neuroeconomics, Behavioral Economics and The Political Economy of Nudge, p. 37-38. Available at: https://www.researchgate.net/publication/319942697_Neuroeconomics_Behavioral_Economics_and_The_Political_Economy_of_Nudge

evolved equally from fields such as *user experience design*[147], *affective computing*[148] and *persuasive computing*[149] on the one hand, and the technologies and infrastructures that have been developed for *online behavioral advertising*[150] on the other. While the former have long been focusing on environments that certain vendors directly control themselves, the latter expands persuasive practices to platforms and services controlled by others. This includes not only the delivery of targeted digital ads based on someone's web searches, browser history, or app usage throughout the digital world, but also, in reverse, the utilization of personal data from many different sources within a company's own environments.

To influence behaviors with data-driven persuasion techniques, businesses and other institutions seem to accept a lower accuracy of the data[151] than they do for systems that make single consequential decisions about people – the only relevant standard is that the data and its application helps them better reach their overall goals, whether those be economic, political or otherwise. Minimal improvements such as small increases in revenue or a certain change in voter turnout are often considered successes.[152] In other cases, however, data might be quite accurate, and data companies constantly aim to improve data quality.[153] Either way, the permanent profiling, evaluation, sorting, and ranking of people according to an organization's goals[154] is the basis for the persuasive practices examined in this section. Any further intervention on people benefits, for instance, from being able to calculate the probable "customer lifetime value"[155] or "return of investment" for every interaction with a person in real-time. As such, data-driven persuasion strategies may also include personalized pricing, based on both business objectives and knowledge about the targets.

It is certainly often not easy to distinguish between practices such as informing, nudging, influencing, manipulation, deception, or even data-driven "coercion". Of course there are many ways in which personalization and tailored user experiences that consider psychological and

[147] See e.g. Egger, F.N. (2001). Affective Design of E-Commerce User Interfaces: How to Maximise Perceived Trustworthiness. In: Helander, M., Khalid, H.M. & Tham (Eds.), Proceedings of CAHD2001: Conference on Affective Human Factors Design, Singapore, June 27-29, 2001: 317-324. Available at: http://www.webusability.ch/articles/CAHD2001.htm

[148] See e.g. Armony, Jorge L (1997): Affective Computing. Trends in Cognitive Sciences , Volume 2 , Issue 7 , 270

[149] Fogg, B. J. (2002): Persuasive technology: using computers to change what we think and do. Ubiquity 2002.

[150] See e.g. Yan, Jun; Ning Liu, Gang Wang, Wen Zhang, Yun Jiang, and Zheng Chen (2009): How much can behavioral targeting help online advertising?. In Proceedings of the 18th international conference on World wide web (WWW '09). ACM, New York, NY, USA, 261-270. Available at: http://dl.acm.org/citation.cfm?id=1526745

[151] See e.g. Schiff, Allison (2017): More Than Half Of Age Data In Mobile Exchanges Is Inaccurate. AdExchanger, January 11, 2017. Available at: https://adexchanger.com/data-exchanges/half-age-data-mobile-exchanges-inaccurate/

[152] For example, when 0.35 % of users click on an online ad, an increase of 0.05% may already considered a success. For an overview of conversion rates in online marketing see e.g. https://databox.com/industry-specific-marketing-kpi-benchmarks

[153] For example, by systematically aggregating identifiers in order to link and combine profiles in more reliable ways, see e.g. Stanhope, Joe; Mary Pilecki; Fatemeh Khatibloo; Tina Moffett; Arleen Chien; Laura Glazer (2016): The Strategic Role Of Identity Resolution. Identity Is Context In The Age Of The Customer. Forrester, October 17, 2016.

[154] See previous section.

[155] In marketing, the "customer lifetime value" is a prediction of the net profit attributed to the entire future relationship with a customer. See e.g. Abdolvand, Neda; Amir Albadvi; Hamidreza Koosha (2014): Customer Lifetime Value: Literature Scoping Map, and an Agenda for Future Research. International Journal of Management Perspective, Vol. 1, No.3, pp. 41-59.

emotional aspects can have enormous benefits for people and be used for good. However, in many cases, they are not.

Insecure teenagers and mood experiments. In 2017, a leaked internal Facebook document revealed how the platform provides an advertiser the opportunity to target 6.4 million young Australians in "moments when young people need a confidence boost" such as when they felt "worthless", "insecure", "stressed", "defeated", "anxious", or like a "failure"[156], based on "internal Facebook data" such as posts and photos.[157] Facebook claimed that this research "was never used to target ads"[158]. However, the company is building a neuroscience lab[159], and is openly promoting research about neuroscience[160], as well on how marketers may "capitalize" on "very important, highly personal and uniquely relevant moments" of users.[161] Back in 2012, Facebook ran its notorious "mood experiment" on nearly 700,000 users that involved manipulating the amount of emotionally positive and negative posts in the users' news feeds, which in turn ended up demonstrably influencing how many emotionally positive and negative messages the users posted themselves. The result of this study, conducted without the users' knowledge, was later published as a research paper that claims to provide "experimental evidence that emotional contagion occurs".[162] Similarly, the ride-hailing platform Uber has not only been accused of abusing its data power to manipulate both its drivers and riders,[163] but also to identify, block, and undermine regulators[164], suppliers,[165] and rivals[166].

Data power and information asymmetries. As Ryan Calo and Alex Rosenblat suggest[167], today's tech companies may be "leveraging their access to information about users and their control over the user experience to mislead, coerce, or otherwise disadvantage" them. For example, they may "reach consumers at their most vulnerable, nudge them into overconsumption, and charge each consumer the most he or she may be willing to pay".[168] Generally, firms can increasingly "use what they know about consumers" to not only "match them to content they might prefer" but also to "nudge consumers to pay more, to work for less, and to behave in

[156] Tiku, Nitasha (2017): Get Ready for the Next Big Privacy Backlash Against Facebook. Wired, 05.21.17. Available at: https://www.wired.com/2017/05/welcome-next-phase-facebook-backlash/

[157] Levin, Sam (2017): Facebook told advertisers it can identify teens feeling 'insecure' and 'worthless'. Guardian, May 1, 2017. Available at: https://www.theguardian.com/technology/2017/may/01/facebook-advertising-data-insecure-teens

[158] Tiku (2017)

[159] Swant, Marty (2017): Facebook Is Building Its Own Neuroscience Center to Study Marketing. Adweek, May 23, 2017. Available at: http://www.adweek.com/digital/facebook-is-building-its-own-neuroscience-center-to-study-marketing/

[160] https://www.facebook.com/iq/articles/mobile-minded-the-small-screen-isnt-so-small

[161] https://fbinsights.files.wordpress.com/2015/09/facebookiq_moments_whitepaper.pdf

[162] Kramer, Adam D. I.; Jamie E. Guillory; Jeffrey T. Hancock (2014): Experimental evidence of massive-scale emotional contagion through social networks. PNAS vol. 111 no. 24, 8788–8790. Available at: http://www.pnas.org/content/111/24/8788.full

[163] http://boingboing.net/2017/03/09/weaponized-information-asymmet.html

[164] https://www.nytimes.com/2017/03/03/technology/uber-greyball-program-evade-authorities.html

[165] https://mobile.nytimes.com/2017/04/23/technology/travis-kalanick-pushes-uber-and-himself-to-the-precipice.html

[166] https://www.theguardian.com/technology/2017/apr/13/uber-allegedly-used-secret-program-to-cripple-rival-lyft

[167] Calo, Ryan and Rosenblat, Alex (2017): The Taking Economy: Uber, Information, and Power (March 9, 2017). Columbia Law Review, Vol. 117, 2017; University of Washington School of Law Research Paper No. 2017-08. Available at: https://ssrn.com/abstract=2929643

[168] Ibid.

other ways that advantage a firm".[169] This encapsulates one of the key concerns that scholars, privacy activists and consumer rights advocates have been raising for years. When a rapidly growing number of daily interactions and behaviors undergo unrestricted digital monitoring, analysis, and assessment, corporate actors can systematically abuse their resultant unprecedented data wealth for their economic advantage. Omer Tene and Jules Polonetsky compare the relationship between data companies and individuals to a "game of poker where one of the players has his hand open and the other keeps his cards close".[170]

Persuasive marketing, instant personalization and testing

Whether based on personal data or not, there are several ways companies can utilize their powerful position within digital environments. Uber, for instance, displayed car icons to consumers ordering a ride on its app's map, suggesting the presence of an available car near them, when, in fact, no car was present at the place shown. After a public controversy, Uber claimed that these car icons were "more of a visual effect" and not meant to mislead users.[171] Such misleading user interface design practices have been referred to as "dark patterns".[172] A guide on "gamification" by the software and data company Oracle has several recommendations for companies on how to exploit peoples' cognitive biases.[173] Examples include suggesting the use of countdowns to "give users some sense of urgency", taking advantage of people's "loss aversion" because of their "psychological tendency to evaluate potential losses as larger and more significant than equivalent gains", or getting "users to act by suggesting that something is available for only a limited time".[174]

Personalized persuasion. Of course, marketing has always been about persuading consumers[175] and while practices that exploit cognitive biases are problematic, they are not what today's personalized data-driven persuasion is about. Marketers have been exploiting personal weaknesses and biases for decades; mundane examples include pricing a product €9.99 rather than €10, as the former is perceived as being closer to €9 than to €10,[176] or placing sweets at eye level of young children. However, as Ryan Calo writes in his study of "digital market manipulation", the "digitization of commerce dramatically alters the capacity of firms to influence consumers at a personal level".[177] When a company "can design an environment from scratch,

[169] Ibid.

[170] Tene, Omer and Jules Polonetsky (2013): Big Data for All: Privacy and User Control in the Age of Analytics. 11 Nw. J. Tech. & Intell. Prop. 239 (2013).p. 255. Available at: http://scholarlycommons.law.northwestern.edu/njtip/vol11/iss5/1

[171] Calo, Ryan and Rosenblat, Alex (2017), p. 28-29

[172] See e.g. Singer, Natasha (2016): When Websites Won't Take No for an Answer. New York Times, May 14, 2016. Available at: https://www.nytimes.com/2016/05/15/technology/personaltech/when-websites-wont-take-no-for-an-answer.html

[173] Oracle Gamification Guidelines. Available at: http://www.oracle.com/webfolder/ux/Applications/uxd/assets/sites/gamification/index.html [20.07.2016]

[174] Ibid.

[175] See e.g. Packard, Vance (rev. ed. 1981): The Hidden Persuaders

[176] Hanson, Jon D. & Douglas A. Kysar (1999): Taking Behavioralism Seriously: The Problem of Market Manipulation, 74 N.Y.U. L. REV. 630 (1999).

[177] Calo, Ryan (2013)

track consumer behavior in that environment, and change the conditions throughout that environment based on what the firm observes, the possibilities to manipulate are legion".[178] Today's companies have access to extensive digital profiles, including someone's financial situation, behaviors, movements, daily routine, search terms, social relationships, interests, and weaknesses. They can choose the right persuasion strategies with the right message at the right time for the right person, monitor and analyze how said individual reacts, and then continuously adapt how they are addressed.

Instant personalization. Personalization based on rich profile information and pervasive real-time monitoring has become a powerful toolset to influence people's behavior in the digital world to, for example, make consumers visit a website, click on an ad, register for a service, subscribe to a newsletter, download an app, or purchase a product. Direct marketing has long been working on personalizing direct mail, call center, and email communication, among other aspects of customer treatment.[179] Companies learned to use information and data mining techniques to identify, acquire, and retain profitable customers, to calculate a person's future "customer lifetime value"[180], to "effectively allocate resources" only to the "most profitable group of customers",[181] and to prevent customer attrition.[182] They learned how to include or exclude groups of customers from certain efforts, how to better influence their behaviors based on data, and how to measure and optimize the outcomes.[183] Now, companies can do personalization in real-time, across devices and communication channels. Data can now be used not only to display ads on websites or within mobile apps, but also on a company's own website, to dynamically personalize the contents, options, and choices offered to both known customers and seemingly "anonymous" visitors. For example, online stores can personalize how they address someone, which products they display prominently, and even the prices of products or services on an individual basis.[184]

Experimenting on users. To further improve this, companies have started continuously experimenting on people. They conduct tests with different variations of functionalities, website designs, user interface elements, headlines, button texts, images, or even different discounts and prices, and then carefully monitor and measure how different groups of users interact with these variations. News organizations, including large outlets such as the Washington

[178] Calo, Ryan and Rosenblat, Alex (2017): The Taking Economy: Uber, Information, and Power (March 9, 2017). Columbia Law Review, Vol. 117, 2017; University of Washington School of Law Research Paper No. 2017-08. Available at: https://ssrn.com/abstract=2929643

[179] Christl (2017), p. 40-53, 75-77

[180] Christl and Spiekermann (2016), p.79

[181] Ngai, E. W. T., Li Xiu, and D. C. K. Chau (2009): Review: Application of data mining techniques in customer relationship management: A literature review and classification. Expert Systems with Applications 36, 2 (March 2009), 2592-2602. http://dx.doi.org/10.1016/j.eswa.2008.02.021

[182] Christl and Spiekermann (2016), p. 127-128

[183] Christl and Spiekermann (2016), p. 125; 127-129

[184] Christl (2017), p. 40-53, 75-77

Post[185], also use such mechanisms with different versions of article headlines in order to figure out which variation performs better.[186] Optimizely, a major technology provider for automated testing that can be used by marketers, news organizations, and any other company providing digital services, offers its clients the ability to "experiment broadly across the entire customer experience, on any channel, any device, and any application".[187]

Exploiting cognitive biases and data to influence behavior

Many of the influencing strategies used in today's marketing stem from neuroeconomics and behavioral economics, which both "share a common interest in the cognitive, affective, and behavioral responses of humans to information about the myriad choices that confront them throughout their lives".[188] As Oscar Gandy has summarized, behavioral scientists have found that adults have limited abilities to "allocate their attention to more than a comparatively small number of relevant features in their environment", "recall facts and experiences", "organize comparisons and evaluations in a consistent manner", or "assign probabilities and appropriate weights to a variety of threats and opportunities".[189] Other examples of the limitations to make rational decisions include the "present bias" that lets people overvalue immediate rewards at the expense of long-term intentions, the "optimism bias" that makes people believe they are at a lower risk for experiencing a negative event than others are, and the "anchoring bias" that lets people rely too much on the first piece of information seen.[190] Additionally, people react differently to information depending on how it is "framed"; as such, a loss may be framed as a gain and a surcharge or tax as a discount.[191]

Using data to influence behavior. Some of these biases may be utilized on digital platforms to influence behaviors without personalizing them to certain individuals; others may become more effective when they are based on personal data; yet others do not make any sense without feeding them with profile information. The much-discussed "confirmation bias", which seems to be the reason for the famous "filter bubble"[192], leads us to "adjust new information in ways that make it easier to accommodate within our already existing set of beliefs".[193] It is obvious that companies that want to influence behavior may utilize personal information about people

[185] https://www.wsj.com/articles/washington-posts-bandit-tool-optimizes-content-for-clicks-1454960088
[186] https://blog.upworthy.com/why-the-title-matters-more-than-the-talk-867d08b75c3b
[187] https://www.optimizely.com [03.05.2017]
[188] Gandy, Jr, Oscar (2017): Neuroeconomics, Behavioral Economics and The Political Economy of Nudge, p. 37-38. Available at: https://www.researchgate.net/publication/319942697_Neuroeconomics_Behavioral_Economics_and_The_Political_Economy_of_Nudge
[189] Ibid.
[190] Hanson, Jon D. and Douglas A. Kysar (1999): Taking Behavioralism Seriously: Some Evidence of Market Manipulation, 112 HARV. L. REV. 1420, 1564–65
[191] Gandy, Jr, Oscar (2017), p. 17
[192] Eli Pariser: The Filter Bubble: What the Internet Is Hiding from You. Penguin Press, New York, 2011
[193] Gandy, Jr, Oscar (2017), p. 18

to present them with messages that reflect their beliefs. This is, of course, used in commercial marketing, but also in election campaigns.[194]

Maurits Kaptein et al point to the concept of **persuasion profiles** as “sets of estimates on the effectiveness of particular influence-strategies on individuals, based on their past responses to these strategies”.[195] As an example, they describe how digital marketers may try to analyze whether someone is more likely to be influenced by what “other people” do, or by what “authorities” do. They then select the most effective strategy for the target and may additionally transfer the extracted “persuasion profiles” to other digital platforms.[196]

Some examples. Many companies in online marketing and advertising technology provide services that help clients optimize their data-driven influence strategies. Some of them specialize in specific use cases. For example, the UK-based firm Nudgr “monitors the behaviour of potential customers”[197] on a website and promises to “automatically engage visitors who will leave without buying”[198] by triggering “perfectly timed” discounts, incentives, or “social proof messages”.[199] The data analytics company CognitiveScale claims to build “cognitive profiles”[200] for engaging shoppers through personalized recommendations “at the right time, right place and with clear evidence to help nudge them towards a desired behavior”.[201] Profiles can be based on “pricing, color, fit, style preferences, digital engagement patterns, and prior shopping history”.[202] The company provides solutions for commerce, but also for financial services[203] and healthcare. For example, they build patient profiles based on “medication, diet, schedule, lifestyle, and socio-economic needs”.[204]

The US-based firm Motimatic offers an “automatic motivational support system”[205] to “deliver highly targeted messages that drive economically beneficial behavior” for clients in education, financial services, insurance, and healthcare.[206] Blending the “latest advances in online advertising technology and motivation science”,[207] their system offers to “drive personal behavior

[194] See e.g. https://medium.com/@privacyint/cambridge-analytica-explained-data-and-elections-6d4e06549491

[195] Kaptein, Maurits; Dean Eckles, and Janet Davis (2011): Envisioning Persuasion Profiles: Challenges for Public Policy and Ethical Practice. 9/10 Interactions 66-69, 66; Available at: https://www.semanticscholar.org/paper/Envisioning-persuasion-profiles-challenges-for-KapteinEckles/fe5f2029df491bdea2cf46697b2e4145c1e226f2/pdf

[196] Ibid.

[197] https://nudgr.io/ [25.09.2017]

[198] https://nudgr.io/about_nudgr [25.09.2017]

[199] https://nudgr.io/ [25.09.2017]

[200] https://www.cognitivescale.com/solutions/commerce/ [30.09.2017]

[201] https://www.cognitivescale.com/products/ [30.09.2017]

[202] https://www.cognitivescale.com/solutions/commerce/ [30.09.2017]

[203] https://www.cognitivescale.com/solutions/financial-services/ [30.09.2017]

[204] https://www.cognitivescale.com/solutions/healthcare/ [30.09.2017]

[205] http://www.motimatic.com/ [30.09.2017]

[206] http://www.motimatic.com/about-us/ [30.09.2017]

[207] Ibid.

through subtle psychological cues, habit forming triggers and other motivational strategies", based on concepts such as "scarcity", "authority", "social proof", "hot triggers", or "nudges".[208]

Predictive marketing and personalized pricing

On a larger scale, many companies provide so-called "predictive marketing" services[209] that combine different aspects of marketing and advertising technology, data, analytics, and personalization. RocketFuel, for example, claims to have "2.7 billion unique profiles" in its data store[210] and offers clients the ability to "bring together trillions of digital and real-world signals to create individual profiles and deliver personalized, always-on, always-relevant experiences to the consumer".[211] The company says that it "scores every impression for its propensity to influence the consumer".[212] Krux, a data management service owned by the customer data giant Salesforce, explains that by monitoring consumers and getting a "granular understanding of individual interests and behaviors", companies can "influence high-potential prospects exactly when they're ready to engage".[213] Twitter's predictive marketing platform TellApart promises to calculate a "customer value score" for each shopper and product combination, a "compilation of likelihood to purchase, predicted order size, and lifetime value" and helps assemble pieces such as "product imagery, logos, offers and any metadata" into personalized content for ads, emails, and websites.[214]

Personalized pricing. Similar methods can be used to personalize prices in online shops by, for example, making predictions as to how valuable someone might be as customer in the long-term or how much someone may be willing to pay at that moment. Strong evidence suggests that online shops already show differently priced products to different consumers, or even different prices for the same products, based on personal characteristics and past behaviors.[215] Of course, dynamic pricing, without being personalized, has been a common practice for a long time – from supermarkets to travel booking. Prices vary depending on the time of a purchase or booking, inventory, available seats, popularity of a product, or prices of competitors. It is also usual to customize pricing depending on the number of units bought – and based on very specific attributes of consumers, for example with discounts for children, families, or elders.[216]

[208] http://www.motimatic.com/ [30.09.2017]

[209] See e.g. https://martechtoday.com/library/predictive-marketing

[210] http://rocketfuel.com/wp-content/uploads/DSP_2015.pdf [03.05.2017]

[211] https://rocketfuel.com/predictive-marketing [03.05.2017]

[212] http://rocketfuel.com/wp-content/uploads/DSP_2015.pdf [03.05.2017]

[213] http://www.krux.com/data-management-platform-solutions/customer-journeys [14.05.2017]

[214] https://www.tellapart.com/platform/ [03.05.2017]

[215] Christl and Spiekermann (2016), p. 41

[216] Borgesius; Frederik J. Zuiderveen (2015): Online Price Discrimination and Data Protection Law (August 28, 2015). Forthcoming as a conference paper for the Amsterdam Privacy Conference 23-26 October 2015; Amsterdam Law School Research Paper No. 2015-32; Institute for Information Law Research Paper No. 2015-02. Available at: http://ssrn.com/abstract=2652665

What is new is that it is now possible to personalize pricing in real-time, based on digital profiles about characteristics or behaviors of consumers. However, since companies such as Amazon already vary their prices up to 2.5 million times on an average day[217], it is difficult to prove whether and, if so, to what extent companies incorporate user characteristics or behavior into their dynamic pricing. Individualized pricing based on personal information further undermines the consumers' ability to compare prices and to break through the opacity of the underlying algorithms, as well as the interests that feed into them. It leads to the same uncertainty and Kafkaesque experiences that the previous section examined. Consumers cannot know whether they see personalized prices or selections of differently priced offers at all, and if so, in what ways their data about past behaviors or other personal properties influence them. Was it because someone acted in a specific way before? Was it because someone visited a specific website, used a specific mobile app, bought a specific product in the supermarket, or watched a specific TV program?

Data-driven coercion? The ride-sharing platform Uber uses a system that automatically increases prices when demand is very high, such as during rush hour, sporting events, or terrorist attacks.[218] While believers in a fully efficient marketplace with rational actors might defend such practices, the company itself has revealed that their data scientists have also studied in which other situations consumers might be willing to pay more. Uber found that users would be ready to pay more when their phone battery is low. The company claims that it currently does not utilize this information.[219] We do not know how likely it is that companies who are in a position to make assessments about a consumer's vulnerabilities across a broad range of life situations might take advantage of this. If they did, such practices should be considered as "data-driven coercion" rather than "personalized pricing". A similar example would be a flight booking platform that increases the prices shown when someone is looking for the same trip several times or when it knows that someone had already booked a hotel room for a certain destination and date range.

In 2017, Uber introduced "route-based pricing", which, according to a company representative indirectly cited by Bloomberg, calculates the "propensity for paying a higher price for a particular route at a certain time of day".[220] For example, someone "traveling from a wealthy neighborhood to another tony spot might be asked to pay more than another person heading to a

[217] https://www.digitalcommerce360.com/2014/12/30/right-price-not-lowest-price/

[218] Vinik, Danny (2014): Uber's Prices Surged in Sydney During the Hostage Crisis, and Everyone Is Furious. New Republic, December 15, 2014. Available at: https://newrepublic.com/article/120564/during-terrorist-attack-sydney-uber-imposing-surge-pricing

[219] See e.g. Calo, Ryan and Rosenblat, Alex (2017): The Taking Economy: Uber, Information, and Power (March 9, 2017). Columbia Law Review, Vol. 117, 2017; University of Washington School of Law Research Paper No. 2017-08. Available at: https://ssrn.com/abstract=2929643

[220] Newcomer, Eric (2017): Uber Starts Charging What It Thinks You're Willing to Pay. Bloomberg, May 19, 2017. Available at: https://www.bloomberg.com/news/articles/2017-05-19/uber-s-future-may-rely-on-predicting-how-much-you-re-willing-to-pay

poorer part of town, even if demand, traffic, and distance are the same".[221] While this might sound fair on the surface, it of course equally affects people traveling from a "poorer part of the town" to the "wealthy neighborhood" and back for work. While this mechanism differs from one that exploits information about personal vulnerabilities such as a low battery level, it demonstrates digital platforms' readiness to take advantage of their data power in order to optimize their bottom line.

Election campaigns and voter targeting

While the commercial sphere exploits cognitive biases such as the "confirmation bias" to great effect in personalizing communication, such tactics also perfectly fit the needs of election campaigns. The practice of targeting voters with personalized messages adapted to their personality and political views on certain issues has already raised massive debates about the potential for political manipulation.[222] A modified form of this process can also be applied in order to, for example, unsettle and discourage people from voting for a candidate by targeting them with digital messages that sow the seeds of doubt in a way targeted to their particular profiles.

According to a Trump campaign official, the US presidential campaign in 2016 had "three major voter suppression operations under way" that targeted white liberals, young women, and African-Americans with Facebook posts that portrayed Hillary Clinton as racist[223], thereby making her appear less appealing to these groups. It is not clear to what extent such data-driven targeting contributed to the election result, and data-driven persuasion is certainly not capable of making people believe the opposite of what they believed before. Nevertheless, evidence suggests that voter turnout of can be effectively increased in certain groups of people with only small changes to the content shown in their Facebook newsfeed.[224] The mechanics of today's social media platforms can amplify such effects.[225] As such, data-driven influence strategies can target small groups of people in order to reach a considerable impact; for example, the combination of data and testing can be used to effectively "seed" content to small groups and make it go viral.[226]

The basic ways political campaigns use these technologies bears a close resemblance to their application in commercial marketing. Instead of targeting consumers with a high predicted "customer lifetime value", an election campaign may target likely "undecided" voters or potential supporters, and then automatically factor the return of investment on an individual level

[221] Ibid.

[222] See e.g. https://medium.com/@privacyint/cambridge-analytica-explained-data-and-elections-6d4e06549491

[223] See e.g. Halpern, Sue (2017): How He Used Facebook to Win. The New York Review of Books, June 8, 2017. Available at: http://www.nybooks.com/articles/2017/06/08/how-trump-used-facebook-to-win/

[224] Christl (2017), p.77

[225] Ibid.

[226] Sass, Erik (2013): Study Identifies "Seed Groups" for Spreading Viral Content. MediaPost, September 18, 2013. Available at: https://www.mediapost.com/publications/article/209454/study-identifies-seed-groups-for-spreading-viral.html

by weighting the costs against the probability of influencing their behavior. In recent years, a large ecosystem of companies has emerged that provides technology, data, analytics, and targeting for election campaigns and public affairs.[227] This includes firms that specialize in voter data and polling as well as ones active in both politics and commerce. Of course, similar methods can be used to gain political influence in a variety of contexts other than elections, such as to shape public opinion in support of or against certain issues, or, perhaps, to deepen social polarization.

There have been many debates about data-driven political interventions in the US, UK, France, Spain, and other countries.[228] The associated question of the Russian state's potential involvement in these interventions lies beyond the scope of this paper. However, one thing is clear: today's networks of digital tracking and profiling can clearly be used to systematically influence and manipulate people. Furthermore, while such tactics may be in their infancy now, they will likely become even more effective over the next few years.

Practical use in centralized and decentralized environments

Many companies have access to vast amounts of behavioral data and analytics capabilities and the capacity to conduct large-scale tests and experiments. In this way, they can discover behavioral anomalies, cognitive biases, and other weaknesses and subsequently use this information to sort people into groups based on they respond to being addressed in certain ways.

Four basic components are needed to deploy data-driven persuasion: a direct user relationship, a digital environment, analytics capabilities, and personal data. Major platforms or large companies have everything in place, including detailed user profiles and a reach that allows them to exploit data and user experiments for the development of superior analytics capabilities. They can adapt and tune their environments on several channels, from functionalities to interaction design. Other companies might acquire access to parts of this stack. There are several ways companies may use personal data about individuals in order to influence their behavior:[229]

- A company may use its own user relationship, environment, analytics, and data. Examples of such approaches include platforms such as Facebook, Google, Apple, and Amazon, a multinational that provides several services, or a smaller company that offers a website, app, or other service.

[227] For a good overview see: https://www.linkedin.com/pulse/how-targeted-digital-ad-gets-made-politics-public-jordan-lieberman/

[228] See e.g. Alandete, David (2017): Russian meddling machine sets sights on Catalonia. The global network that acted in favor of Donald Trump and Brexit turns attention to Spain. El Pais, 28 SEP 2017. Available at: https://elpais.com/elpais/2017/09/26/inenglish/1506413477_994601.html

[229] For details about the ways how companies can acquire, link, combine or integrate data and address people across the digital world see: Christl (2017), p. 40-53

- A company may use its own user relationship and environment, but access profile data and/or analytics from a third party. It could acquire additional real-time information about its website or app users and then use this to personalize their individual environments.
- A company may fully utilize the user relationships, environments, analytics, and data of others by, for example, leveraging the capabilities of the large platforms or the programmatic advertising ecosystem.
- Of course, much more complex setups that dynamically combine a company's own user relationships, environments, analytics capabilities, and data with those of others are also possible.

There are a few basic different ways a company may utilize user relationships, environments, analytics, and data from third parties for data-driven persuasion:[230]

- A company may want to discover and target new "unknown" people who are fitting their goals and persuasion strategies throughout the digital world, for example via Facebook, Google, or the programmatic advertising ecosystem. For example, it may target persons with a high predicted "customer lifetime value" who are interested in gambling and have recently been searching online for credit-related topics.
- A company may monitor users "touching" their environment the first time, try to recognize them and learn more about them, decide whether they fit the company's goals and persuasion strategies, and then personalize content and options. This may include seemingly "anonymous" website visitors with certain characteristics and behaviors.
- A company might want to learn more about "known" customers, prospects, members, or users in order to act on them accordingly.

These examples show only a few ways that companies can combine their own capabilities with those of others. Today's networks of digital tracking and profiling provide powerful capabilities to systematically monitor and manage consumer behavior through real-time behavioral feeds that can, according to complex sets of rules and instructions, dynamically react to what people do across many life contexts.[231]

Markets for behavioral control?

The marketing giant GroupM estimates in a report that Facebook makes 200 trillion little decisions a day in order to determine which content may be "relevant" to each of its users.[232] Relevance, though, is a two way street; the report suggests that 'relevance' for Facebook and Google "is about economic outcomes for the company as much as it is about quality of user experi-

[230] Ibid.
[231] Christl (2017), p. 49-50
[232] https://groupmp6160223111045.azureedge.net/cmscontent/admin.groupm.com/api/file/2618 [03.05.2017]

ence".[233] The same applies to personalized communication operated or delivered by myriads of other companies, whether they use a large platform's infrastructures or rely on more decentralized networks of digital tracking and profiling. Which kinds of information, functionalities, options, ads, or prices someone sees is determined by calculations aimed at maximizing efficiency and profit, including whether an individual might be 'worth' the effort at a certain moment.

The author and academic Shoshana Zuboff states that we are not only witnessing the rise of "markets for personal data" but also of "markets for behavioral control", "composed of those who sell opportunities to influence behavior for profit and those who purchase such opportunities".[234] She sees the emergence of a "ubiquitous networked institutional regime that records, modifies, and commodifies everyday experience from toasters to bodies, communication to thought".[235]

3.3 Implications for individuals, groups of people, and society

This section summarizes the technologies and data-driven practices that have been examined in the previous sections and further discusses their social, economic, and political implications.

Companies increasingly and unilaterally shape the networked environments and experiences of everyday life. In the age of digital technology, several factors have contributed to a development that generally weakens the position of individuals – whether as consumers or citizens – against powerful commercial parties and other institutions. The extent of corporate control over today's digital environments certainly raises substantial concerns on its own. The deep and wide-ranging access these corporate actors have to personal data from diverse life contexts – an issue that has long been addressed under the umbrella of information privacy – is one of the major factors producing this shift in power.

With the help of advanced analysis technologies and knowledge extracted from large-scale data aggregation, powerful commercial parties directly or indirectly use personal information about individuals – who are in comparatively vulnerable positions – to constantly evaluate, classify, sort, rate, and rank them according to their business objectives. In many areas, data-driven systems decide individuals' choices, opportunities, and life-chances. This furthermore

[233] Ibid.
[234] Zuboff, Shoshana (2015): Big Other: Surveillance Capitalism and the Prospects of an Information Civilization (April 4, 2015). Journal of Information Technology (2015) 30, 75–89, p. 85. Available at: http://ssrn.com/abstract=2594754
[235] Zuboff, Shoshana (2015): Big Other: Surveillance Capitalism and the Prospects of an Information Civilization (April 4, 2015). Journal of Information Technology (2015) 30, 75–89, p. 81. Available at: http://ssrn.com/abstract=2594754

enables companies to exploit information asymmetries, personal weaknesses, and cognitive biases in order to influence behavior at scale.

By their very nature, **data-driven decisions** are discriminatory in the sense that they use personal information and data mining methods in order to distinguish between people of certain groups. Many things can go wrong in such systems. In the moment when such systems make a decision about individuals, they may use inaccurate or outdated data as an input or make inaccurate or biased inferences about them. In many cases, automated decisions are biased against certain groups of people without this being intended by a company's business objectives. However, such technologies also make it easier to revive traditional forms of intentional discrimination because the intentional prejudicial use of certain categories – such as gender and ethnicity – is easy to hide and difficult to prove. On a more fundamental level, systems that make automated decisions on the basis of personal data tend to reflect not only existing societal inequalities and prejudicial biases, but also the economic goals and business needs of companies. As such, they are almost necessarily biased towards objectives and values such as efficiency and profit maximization. Generally, access to personal information significantly empowers companies and other institutions by providing them with better abilities to "accurately" discriminate between people in order to take advantage of them.

Whether automated decisions based on personal information are technically accurate or inaccurate, intentionally or non-intentionally biased, they may be used in all cases to unfairly or unjustly exclude or target people and thus affect their opportunities and life-chances. These effects may be distributed unevenly across different population groups and can be cumulative over time. Individuals who are subject to consequential data-driven decisions may get excluded from certain opportunities, receive disadvantageous terms and prices, become subject to further investigations, or be filtered out in advance. Being rated and judged as an unwanted, suspicious, or non-valuable person may also lead to being excluded or disadvantaged in ways where the consequences of a single decision are minimal, but consistently affect certain groups of people more than others, and thus produce significant disadvantages on an aggregate societal level.

When data-driven decisions systematically discriminate against already disadvantaged groups, this may not only reproduce, but also increase social and economic inequality at scale. In the worst case, discriminatory predictive systems such as credit scoring become self-fulfilling prophecies, creating the future they pretend to predict. Furthermore, the refusal to share personal information or participate in today's digital tracking may have consequences, too. If not enough data about a person is available, the risk of a customer relationship may be considered as too high by default.

Opacity, chilling effects and loss of autonomy. Many of today's systems making data-driven decisions are opaque and nontransparent, utilize personal data from a wide range of unknown sources, and often operate without the fully informed consent of the subjects affected by the decision; in some cases, they function completely invisibly. Individuals can hardly object to such decisions. For the most part, companies do not explain them. They are rarely held accountable or liable. Decisions that are not fully transparent are, on an individual level, perceived as arbitrary per se. This leads to uncertainty, Kafkaesque experiences, and can produce chilling effects on forms of action or expression, thus limiting individual autonomy.

Even credit scores, which have long been regulated in some way, evince massive deficiencies in regards to accuracy, arbitrariness, and opacity. If systems that affect vital needs such as banking, insurance, employment, or healthcare expand their use of personal information to behavioral or relational data about everyday life – such as browser histories, movements, social media data, and friends lists – this would introduce new forms and increase the extent of digital behavioral control, massively restricting autonomy and fundamentally violating human dignity. Even worse, such far-reaching and privacy-invasive uses of personal data would most likely disproportionally affect already disadvantaged groups.

Of course, a fundamental factor that determines the degree to which such automated decisions that base on personal information affect autonomy is the degree of choice someone has in whether to be subject to them. However, in many cases, individuals do not have a real choice, because refusing to agree would lead to serious social or economic disadvantages. Again, this lack of real choice particularly affects economically or otherwise disadvantaged groups.

As both a subset and expansion of automated decision-making, commercial parties and other institutions can use **personalization and data-driven persuasion** against people, both in environments they themselves control and across the wider digital world. When companies exploit information asymmetries, personal weaknesses, or cognitive biases in order to confine or frame the choices someone has, or to selectively influence behavior, they limit the negotiating power, personal and political agency, autonomy, and dignity of the affected individuals. Based on their access to private information about individuals' lives and behaviors, companies can easily personalize manipulative, misleading, deceptive, or even coercive practices. Such strategies are used in diverse areas such as behavioral advertising, marketing and sales, as well as news, entertainment, and political campaigning. Data-driven persuasion may of course be used not only against adults, but also against children. Not in the least, when such practices systematically target already disadvantaged social groups, they produce ramifications for equality on a societal level.

Manipulative and exploitative uses of personalization capitalize on knowledge generated from large data sets, insights from behavioral science, and from permanent experimentation with real people, a process that typically occurs without the subjects' knowledge. Most importantly, these practices base on access to personal information at the moment of application, which allows companies to sort, rank, include or exclude people, and then customize the digital environments of their targets accordingly. On the basis of this total surveillance of all interactions, the contents, functionalities, and choices on websites, apps, and services can be further adapted in real-time on an individual level. As such, today's distributed networks of digital tracking and profiling – with their extensive real-time data sharing and matching capabilities – have massively expanded the ways in which companies can use personalized persuasion across different platforms, services, devices, and life contexts.

Combined with influencing strategies derived from neuroeconomics and behavioral economics, these capabilities fundamentally undermine the concept of rational choice, and thus the basic foundation of market economy. When used in political campaigns, such data-driven persuasion may undermine democracy at large.

4. Conclusion

Today, companies aggregate and utilize personal information at an unprecedented scale. Powerful commercial parties have seized control of data pertaining to billions of people and built a pervasive, complex, dynamic, and opaque infrastructure that allows them – together with a wide array of other businesses – to constantly monitor, follow, sort, rate, and rank people as they see fit. Today's commercial networks of digital tracking and profiling work by interlinking heterogeneous actors in a decentralized manner; nevertheless, large players structurally dominate the ecosystem. In this way, they increasingly and unilaterally shape the networked environments and experiences that underlie and determine everyday life. Their services allow other companies to plug into this ecosystem and likewise take advantage of extensive personal information.

This working paper shows how the corporate use of personal data can affect individuals, groups of people, and society at large, particularly in the context of automated decisions and data-driven personalization. Systems that make automated decisions about people based on their data produce substantial adverse effects. They are largely opaque, nontransparent, arbitrary, biased, unfair, and unaccountable – even in areas, such as credit scoring, that have long been regulated in some way. Through data-driven personalization, companies and other institutions can easily utilize information asymmetries in order to exploit personal weaknesses with calculated efficiency. Personalized persuasion strategies provide the means to effectively influence behavior at scale; manipulative, misleading, deceptive, or even coercive strategies can be automated and customized down to the individual level.

Based on the examination of business practices and their implications, both in this paper and in preceding research[236], we conclude that, in their current state, today's corporate networks of digital tracking and profiling show a massive potential to limit personal agency, autonomy, and human dignity. This is not only a problem for individuals, but one that affects society at

[236] Christl (2017)

large. By improving the ability to exclude or precisely target already-disadvantaged groups, they inherently tend toward disproportionally affecting these groups and therefore increase social and economic inequality. Combined with influencing strategies derived from neuroeconomics and behavioral economics, data-driven persuasion undermines the concept of rational choice, and thus the basic foundation of market economy. When used in political campaigns or in other efforts to shape public policy, it may have a similar effect on democracy.

While this working paper and preceding report do not directly offer solutions, they examine, document, structure, and contextualize today's commercial personal data industries and their implications; further research will build on this basis. Hopefully, they will also encourage and contribute to further work by others, whether by researchers, scholars, journalists, and stakeholders in the fields of civil rights, data protection, and consumer protection. Ultimately, the report aims to inform policymakers and even companies themselves.

Addendum – A few remarks on how to go forward

We urgently need not only a conversation on how to make existing systems making decisions about people fair, accountable, and transparent[237], but also a debate about what kinds of automated decisions, based on which kinds of personal information, we generally want to tolerate. Crucial questions include: Which behaviors should they be allowed to reward and punish? Which input data related to personal characteristics, behaviors, and lives should banks, insurers, healthcare providers, employers, and other parties be allowed to use in such systems? And not least, who controls these systems? Either way, building systems that prioritize objectives and values such as community, fairness, equality and social welfare over efficiency and profit maximization would certainly lead to different results.

The abuse of personalization and data-driven persuasion gives rise to three major challenges. The first involves the development and implementation of updated legal distinctions between acceptable persuasive practices and beneficial personalization on the one hand and unacceptable manipulation and the exploitation of personal weaknesses on the other.[238] Today's legal frameworks – not to mention the mechanisms of their enforcement – do not seem adequately prepared for a situation in which companies can control data, digital environments, and experiences at such extensive levels.[239] The second relates, in a general way, to addressing the increased power imbalances between companies and consumers inherent to these data-driven environments as they currently exist. In this sense, experimentation and tests on una-

[237] See e.g. conferences such as "Conference on Fairness, Accountability, and Transparency" (https://fatconference.org) or "Fairness, Accountability, and Transparency in Machine Learning" (http://www.fatml.org)

[238] See e.g. Helberger, Natali (2016): Profiling and Targeting Consumers in the Internet of Things – A New Challenge for Consumer Law (February 6, 2016). Available: https://ssrn.com/abstract=2728717

[239] See e.g. Calo, Ryan and Rosenblat, Alex (2017): The Taking Economy: Uber, Information, and Power (March 9, 2017). Columbia Law Review, Vol. 117, 2017; University of Washington School of Law Research Paper No. 2017-08. Available at: https://ssrn.com/abstract=2929643

ware users[240] perhaps deserve consideration as a separate issue. The third involves mitigating the pervasive collection, disclosure, trade, and use of personal data that today occurs across companies and largely happens without the subjects' knowledge and expectation.

With regards to privacy and data protection legislation, minor tweaks can have major consequences for corporate digital tracking and profiling. For example, forcing companies to respect a web browser's "Do Not Track"[241] setting – and advocating for such as the default – would probably undercut much of the tracking that pervades today's web. Making it more difficult to use pseudonymous codes and identifiers to constantly link and match digital profiles across companies for purposes other than the provided services would probably disrupt parts of today's "markets of behavioral control".[242] Failing to do so means resigning to the dystopian perspective of a near future in which a vast array of connected devices – each of them feeding a near-constant stream of personal data to unknown commercial parties – will be even more embedded in everyday life than now.

Admittedly, changing the present tendencies is not an easy task. There are several challenges on a fundamental level. One of them is the need to be able to preserve the distinction between **personal data and anonymity**. The latter constitutes a basic foundation of all privacy and data protection legislation, but the access to large amounts of personal data, cross-linking between data sets, as well as through inferences and de-identification based on data analytics undermine it.[243] However, there are ways to mitigate these problems, from outlawing de-identification[244] to fine-tuning the definition and interpretation of personal data[245] to adding use-based privacy regulation[246] without weakening existing protections. The question of "privacy in public" is a related issue.[247] Conversely, while the unilateral and large-scale corporate utilization of personal data that allows companies to judge and single out people on an indi-

[240] See Christl(2017), p. 77-78

[241] See e.g. https://www.eff.org/de/issues/do-not-track

[242] See section 3.2

[243] See e.g. Barocas, S., & Nissenbaum, H. (2014): Big Data's End Run around Anonymity and Consent. In J. Lane, V. Stodden, S. Bender, & H. Nissenbaum (Eds.), Privacy, Big Data, and the Public Good: Frameworks for Engagement (pp. 44-75). Cambridge: Cambridge University Press. Available at: https://doi.org/10.1017/CBO9781107590205.004

[244] Apart from that reliable exceptions for researchers would be crucial, the main argument against banning re-identification is that a ban would be difficult to enforce. However, many criminal delicts such as insider trading or tax evasion are difficult to enforce, as well. See e.g. Ohm, Paul (2009): Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization (August 13, 2009). UCLA Law Review, Vol. 57, p. 1701, 2010; U of Colorado Law Legal Studies Research Paper No. 9-12, p. 1758. Available: https://ssrn.com/abstract=1450006

[245] It makes a difference whether personal data is defined as "any information relating to an identified or identifiable natural person" as in the upcoming EU GDPR or as "data that can be reasonably linked to a particular person, computer, or device" as recently defined by the US FTC (https://www.ftc.gov/news-events/blogs/business-blog/2016/04/keeping-online-advertising-industry). See also e.g. Borgesius, Frederik J. Zuiderveen (2016): Singling Out People Without Knowing Their Names – Behavioural Targeting, Pseudonymous Data, and the New Data Protection Regulation (February 16, 2016). Available at: http://ssrn.com/abstract=2733115

[246] Hoboken, Joris van (2016): From Collection to Use in Privacy Regulation? A Forward Looking Comparison of European and U.S. Frameworks for Personal Data Processing, In: Van Der Sloot, Broeders and Schrijvers (eds.), Exploring the Boundaries of Big Data, Netherlands Scientific Council for Government Policy, 2016, pp. 231-259.

[247] See e.g. Calo, Ryan (2017): Artificial Intelligence Policy: A Roadmap (August 8, 2017), p. 18. Available at: https://ssrn.com/abstract=3015350

vidual level raises massive concerns, the access to knowledge extracted from "big data" leads to inverse problems. Only a few large companies have access to really comprehensive amounts of data[248], thus creating a "big data divide".[249] Perhaps the measures mentioned above, when combined with technologies such as differential privacy[250], could pave a way towards pushing for broader public access to knowledge aggregated by the "big data rich"[251] *without* making everybody an easy target for commercial data exploitation.

A second basic challenge for privacy legislation is the problem with **consent and choice**. Today, myriads of companies collect vast amounts of personal information about individuals without their *effectively* informed consent and knowledge, although pretending otherwise at a formal level.[252] Better regulating and enforcing the principle of informed consent is certainly crucial in many areas. Technical solutions may help,[253] and although today's privacy policies and terms are often misleading, impossible to understand, and not adequately usable for consumers in their daily routine[254], they will stay essential for the enforcement by data protection authorities, as well for scrutiny through consumer watchdog organizations and others. However, the principles of consent and choice unilaterally shift the responsibility of privacy protection to the individual level,[255] which leads to several problems.

First, it is nearly impossible for consumers to comprehend the mechanisms and possible long-term implications of today's data processing.[256] Second, an issue that is getting increasingly important is that when individuals share data with companies this may also have an impact on the "privacy of others".[257] Third, and most important, refusing to agree to data collection is simply not an option in many cases. Consumers can "hardly avoid privacy contracts" because "almost all banks, software and hardware vendors, social networking sites, digital content services, retail loyalty programs, and telecommunications providers employ them".[258] They can

[248] See section 2.3

[249] Andrejevic, Mark (2014): The Big Data Divide. International Journal of Communication 8 (2014), 1673–1689. Available at: http://ijoc.org/index.php/ijoc/article/download/2161/1163

[250] See e.g. Dwork, Cynthia (2008): Differential privacy: a survey of results. In Proceedings of the 5th international conference on Theory and applications of models of computation (TAMC'08), Manindra Agrawal, Dingzhu Du, Zhenhua Duan, and Angsheng Li (Eds.). Springer-Verlag, Berlin, Heidelberg, 1-19. Available at: http://dl.acm.org/citation.cfm?id=1791836

[251] boyd danah; Crawford, Kate (2012): Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon, Information, Communication & Society 15:5, pp. 662-679. Available at: http://www.tandfonline.com/doi/abs/10.1080/.VB8Tz_l_uCk

[252] See e.g. Christl and Spiekermann (2016), p. 121-123

[253] E.g. standarized icons giving consumers a meaningful overview of data processing, standardized privacy exchange protocols, tools supporting semi-automated privacy self-management, so-called "privacy agents"; see Christl and Spiekermann (2016), p. 143; 148

[254] See e.g. Christl and Spiekermann (2016), p. 121-123

[255] See e.g. Hartzog, Woodrow (2017): Privacy and the Dark Side of Control. The Institute of Art and Ideas, Sept 4, 2017. Available at: https://iainews.iai.tv/articles/privacy-the-dark-side-of-control-auid-882

[256] Ibid.

[257] See e.g. Taylor, L., Floridi, L., van der Sloot, B. eds. (2017): Group Privacy: new challenges of data technologies. Dordrecht: Springer, p. 9.

[258] Rhoen, Michiel (2016): Beyond consent: improving data protection through consumer protection law. Internet Policy Review, 5(1), p2. Available at: http://policyreview.info/articles/analysis/beyond-consent-improving-data-protection-throughconsumer-protection-law

rarely avoid consenting to data collection without opting out of much of modern life. in many cases, people are required to consent, either because offers and services that are not based on invasive digital tracking are not available, or because non-participation would lead to serious social or economic disadvantages in life.

Scott Peppet suggests that "those with valuable credentials, clean medical records, and impressive credit scores will want to disclose those traits to receive preferential economic treatment. Others may then find that they must also disclose private information to avoid the negative inferences attached to staying silent".[259] He points to an "unraveling effect" of privacy that will inevitably lead to a "full disclosure future", as long as no corrective measures are introduced.[260] Large insurers already offer programs that promise considerable discounts depending on someone's driving behavior, health-related activities such as the daily steps, grocery purchases, or visits to the fitness studio.[261] These programs are transparent, and although behavioral data is not (yet) used for up-front eligibility decisions or risk-based pricing, they are openly based on behavioral monitoring and reward behavioral change. With its plans to create a comprehensive "social credit" system constantly judging citizens not only based on their credit history, but also on their behavior on social networks, the Chinese government is already paving the way.[262] As long as people consent to pervasive surveillance of their everyday lives, concepts like the notion of "taking back control of data", declaring personal data a property right, or even the promotion of a right to transparently sell one's personal information[263], will completely fail to address the incremental unraveling of privacy as described by Peppet. This set of problems is closely related to issues of data-driven decision-making and perhaps one of the most serious challenges to address, if we do not want to end up in a future society based on pervasive digital social control and ubiquitous monitoring of everyday life, where privacy and autonomy become – if it remains at all – a luxury commodity for the rich.

Generally, the widely unrestricted development of a digital economy that is based on pervasive surveillance and enables massive information and power asymmetries between corporate parties and individuals has led to a situation where minor tweaks will hardly mitigate the societal challenges resulting from it. In the last 10 years, billions of dollars in venture capital have been poured into funding business models based on the unscrupulous mass exploitation of data, without considering any ethical, societal, cultural, and political implications. Moreover, the shortfall of privacy regulation in the US and the absence of its enforcement in Europe has actively impeded the emergence of *other* kinds of digital innovation, that is, of practices, tech-

[259] Peppet, Scott R. (2010): Unraveling Privacy: The Personal Prospectus & the Threat of a Full Disclosure Future (August 7, 2010). Northwestern University Law Review, 2011 . Available: https://ssrn.com/abstract=1678634

[260] Ibid.

[261] For a detailed description of how insurance programs involving car telematics and wearables see Christl and Spiekermann (2016), p. 52-68

[262] See e.g. https://citizenlab.ca/2017/01/cashless-society-cached-data-security-considerations-chinese-social-credit-system

[263] See e.g. https://github.com/okffi/mydata

nologies, and business models that preserve autonomy, democracy, social justice, and human dignity. Tech giants and industry groups engage in massive lobbying aiming to actively shaping public policy to advance their position.[264] With regards to consumer data brokers Chris Hoofnagle states that legislation to "rein in these companies has been politically impossible to enact, in part because so many large businesses – and politicians themselves – use information brokers to amass data on people".[265]

Key privacy issues. That being said, with regards to the corporate collection and utilization of personal data the most urgent issues to be addressed include:

- the ubiquitous personal data sharing of website, apps, services, and devices with third parties such as data brokers, advertising technology companies, and analytics firms;[266]
- conversely, the availability of third-party data for companies in diverse industries and its use for automated differential treatment of consumers at the individual level;[267]
- any invasive and de-contextualized use of personal information about everyday life behaviors for judgement, risk assessment or risk-based pricing in essential areas of life such as finance, insurance, education, employment, welfare, or law enforcement;[268]
- any use of data collected for identity verification, risk assessment, credit rating, fraud detection, and network security for different purposes, e.g. marketing and sales;[269]
- the use or disclosure of transactional data in telecom, internet access services, banking, and payment for different purposes than to provide these services;[270]
- the platform and data power of tech giants, under special consideration of their increasingly relevant role as providers of verified identities;[271]
- tech intermediaries aiming to "disrupt" traditional industries that try circumventing regulation and operate in grey legal areas, or whose business plan even includes changing the law,[272] deserve special attention regarding their data practices;[273]

[264] See e.g. Christl and Spiekermann (2016), p. 139, and: Romm, Tony (2015): Tech giants get deeper into D.C. influence game. Politico, 01/21/2015. Available at: http://www.politico.com/story/2015/01/tech-lobby-apple-amazon-facebook-google-114468, and: Byers, Alex (2017): How a telecom-tech alliance wiped out FCC's privacy rules. Politico, 03/31/2017. Available at: http://www.politico.com/story/2017/03/broadband-data-victory-republicans-236760

[265] Hoofnagle, Chris Jay (2016): Federal Trade Commission Privacy Law and Policy - Chapter 6 Online Privacy (February 1, 2016). In: Chris Jay Hoofnagle, Federal Trade Commission Privacy Law and Policy (Cambridge University Press 2016); UC Berkeley Public Law Research Paper No. 2800276. Available: https://ssrn.com/abstract=2800276

[266] Christl and Spiekermann (2016), p. 45-75

[267] See chapter 3

[268] Christl (2017), p. 27-39, 79-81

[269] Christl (2017), p. 79-81

[270] Christl (2017), p. 18-22

[271] See e.g. Christl (2017), p. 11-26, and e.g. http://www.businessinsider.com/facebook-third-party-identity-provider-2015-7

[272] See e.g. Pollman, Elizabeth and Barry, Jordan M. (2017): Regulatory Entrepreneurship (March 3, 2016). 90 S. Cal. L. Rev. 383 (2017); Loyola Law School, Los Angeles Legal Studies Research Paper No. 2017-29. Available at: https://ssrn.com/abstract=2741987

[273] For example, platforms that claim to be intermediaries, but are in fact employers. The risk that such platforms utilize highly invasive kinds of behavioral data about their de-facto employees is higher as traditional employers. See e.g. http://www.huffingtonpost.com/entry/uber-monitor-drivers_us_56aed04ce4b00b033aafa03f

While addressing these issues will be more difficult in the US and other regions with weak legal privacy frameworks, the upcoming new European privacy legislation, which includes both the already adopted EU General Data Protection Regulation (GDPR) and the still disputed ePrivacy Regulation, might ban or at least slow down some of the most irresponsible and invasive practices of third-party data collection.[274] Depending on its final implementation and depending on its practical interpretation and enforcement, it might be a massive step forward. Providing instruments to push for more transparency, accountability, and liability, privacy and consumer advocates can take advantage of this unique opportunity to move forward. Other regulatory instruments such as anti-discrimination, consumer protection, and competition law are equally important in order to challenge unfair discrimination, information asymmetries and power imbalances[275], as well the dominance of certain large players that nobody can escape.[276]

Demanding and enforcing transparency about personal data collection, disclosure, analysis, and use has certainly its limits[277]. It will not always directly empower individuals who already cannot handle the information overload caused by thousands of pages in privacy policies, and it will never be a replacement for solid protections. However, given the extent of opacity and non-transparency currently in place, it empowers individuals indirectly by providing authorities, advocates, journalists, and others with powerful means to address questionable practices and raise awareness. Research, investigation, raising awareness, and legal action are certainly the basis for being able to cope with the market power, resources, and lobbying efforts of today's personal data industries. Single initiatives do have an exceptional impact,[278] as well do coordinated efforts[279] across different kinds of stakeholders, including consumer, digital rights and civil rights organizations, as well as universities, media, privacy and law professionals, data protection authorities, and parts of the industry working on privacy-preserving technologies and business models.

Anyway, without any doubt, in the face of current commercial practices it will require a major collective effort to make a positive vision of a future information society reality.

[274] A study commissioned by the Interactive Advertising Bureau Europe (IAB) expects a considerable impact on data collection by companies that do not have a direct relationship with consumers, see: IHS Markit (2017): The economic value of behavioural targeting in digital advertising. Analysis on behalf of IAB Europe and EDAA. Available at: https://www.iabeurope.eu/wp-content/uploads/2017/09/BehaviouralTargeting_FINAL.pdf

[275] See e.g. Rhoen, M. (2016): Beyond consent: improving data protection through consumer protection law. Internet Policy Review, 5(1). Available at: https://policyreview.info/articles/analysis/beyond-consent-improving-data-protection-through-consumer-protection-law

[276] See e.g. http://www.reuters.com/article/us-facebook-germany-dataprotection-idUSKCN0W40Y7

[277] See e.g. Ananny, Mike and Kate Crawford (2016): Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. New Media & Society. Available at: https://doi.org/10.1177/1461444816676645i, or: Crain, Matthew (2016): The limits of transparency: Data brokers and commodification. New Media & Society. Available at: https://doi.org/10.1177/1461444816657096

[278] See e.g. https://www.propublica.org/people/julia-angwin and https://en.wikipedia.org/wiki/Max_Schrems

[279] See e.g. https://www.forbrukerradet.no/siste-nytt/connected-toys-violate-consumer-laws/

References

Abdolvand, Neda; Amir Albadvi; Hamidreza Koosha (2014): Customer Lifetime Value: Literature Scoping Map, and an Agenda for Future Research. International Journal of Management Perspective, Vol. 1, No.3

Ananny, Mike and Kate Crawford (2016): Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. New Media & Society. Available at: https://doi.org/10.1177/1461444816676645i

Andrejevic, Mark (2014): The Big Data Divide. International Journal of Communication 8 (2014), 1673–1689. Available at: http://ijoc.org/index.php/ijoc/article/download/2161/1163

Angwin et al (2017): Minority Neighborhoods Pay Higher Car Insurance Premiums Than White Areas With the Same Risk. ProPublica, April 5, 2017. Available at: https://www.propublica.org/article/minority-neighborhoods-higher-car-insurance-premiums-white-areas-same-risk

Armony, Jorge L (1997): Affective Computing. Trends in Cognitive Sciences , Volume 2 , Issue 7 , 270

Barocas, S.; Nissenbaum, H. (2014): Big Data's End Run around Anonymity and Consent. In J. Lane, V. Stodden, S. Bender, & H. Nissenbaum (Eds.), Privacy, Big Data, and the Public Good: Frameworks for Engagement (pp. 44-75). Cambridge: Cambridge University Press. Available at: https://doi.org/10.1017/CBO9781107590205.004

Barocas, Solon (2014): Data Mining and the Discourse on Discrimination. Conference on Knowledge Discovery and Data Mining. Available at: https://dataethics.github.io/proceedings/DataMiningandtheDiscourseOnDiscrimination.pdf

Barocas, Solon and Andrew Selbst (2016): Big Data's Disparate Impact, California Law Review, Vol. 104, 2016. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899

Borgesius; Frederik J. Zuiderveen (2015): Online Price Discrimination and Data Protection Law (August 28, 2015). Forthcoming as a conference paper for the Amsterdam Privacy Conference 23-26 October 2015; Amsterdam Law School Research Paper No. 2015-32; Institute for Information Law Research Paper No. 2015-02. Available at: http://ssrn.com/abstract=2652665

boyd danah; Crawford, Kate (2012): Critical Questions for Big Data: Provocations for a Cultural, Technological, and Scholarly Phenomenon, Information, Communication & Society 15:5. Available at: https://doi.org/10.1177/1461444816676645

Calo, Ryan (2013): Digital Market Manipulation (August 15, 2013). 82 George Washington Law Review 995 (2014); University of Washington School of Law Research Paper No. 2013-27. Available at: https://ssrn.com/abstract=2309703

Calo, Ryan (2017): Artificial Intelligence Policy: A Roadmap (August 8, 2017). Available at: https://ssrn.com/abstract=3015350

Calo, Ryan and Rosenblat, Alex (2017): The Taking Economy: Uber, Information, and Power (March 9, 2017). Columbia Law Review, Vol. 117, 2017; University of Washington School of Law Research Paper No. 2017-08. Available at: https://ssrn.com/abstract=2929643

Christl, Wolfie and Sarah Spiekermann (2016): Networks of Control. A Report on Corporate Surveillance, Digital Tracking, Big Data & Privacy. Facultas, Vienna 2016. Available at: http://crackedlabs.org/en/networksofcontrol

Christl, Wolfie (2017): Corporate Surveillance in Everyday Life. How Companies Collect, Combine, Analyze, Trade, and Use Personal Data on Billions. Report by Cracked Labs, June 2017. Available at: http://crackedlabs.org/en/corporate-surveillance

Citron, Danielle Keats and Frank A. Pasquale (2014): The Scored Society: Due Process for Automated Predictions. Washington Law Review, Vol. 89, 2014, U of Maryland Legal Studies Research Paper No. 2014-8. Available at: https://ssrn.com/abstract=2376209

Clarke, Roger (1988): Information Technology and Dataveillance. Commun. ACM 31, 5 (May 1988), 498-512. Available at: http://www.rogerclarke.com/DV/CACM88.html

Constine, Josh (2014): Facebook Stops Irresponsibly Defaulting Privacy Of New Users' Posts To "Public", Changes To "Friends". TechCrunch, May 22, 2014. Available at: https://techcrunch.com/2014/05/22/sometimes-less-open-is-more/

Crain, Matthew (2016): The limits of transparency: Data brokers and commodification. New Media & Society. Available at: https://doi.org/10.1177/1461444816657096

De Zwart, Melissa; Humphreys, Sal; Van Dissel, Beatrix (2014): Surveillance, big data and democracy: lessons for Australia from the US and UK, UNSW Law Journal. Available at: http://www.unswlawjournal.unsw.edu.au/sites/default/files/final_t3_de_zwart_humphreys_and_van_dissel.pdf

DECODE (2017): Me, my data and I: The future of the personal data economy. DECODE report, September 2017. Available at: https://decodeproject.eu/publications/me-my-data-and-ithe-future-personal-data-economy

Dumortier, Franck (2009): Facebook and Risks of "De-contextualization" of Information. In: Monograph "5th Internet, Law and Politics Congress. The Pros and Cons of Social Networks", Universitat Oberta de Catalunya. Available at: https://works.bepress.com/franck_dumortier/1/

Dwork, Cynthia (2008): Differential privacy: a survey of results. In Proceedings of the 5th international conference on Theory and applications of models of computation (TAMC'08), Manindra Agrawal, Dingzhu Du, Zhenhua Duan, and Angsheng Li (Eds.). Springer-Verlag, Berlin, Heidelberg, 1-19. Available at: http://dl.acm.org/citation.cfm?id=1791836

Egger, F.N. (2001): Affective Design of E-Commerce User Interfaces: How to Maximise Perceived Trustworthiness. In: Helander, M., Khalid, H.M. & Tham (Eds.), Proceedings of CAHD2001: Conference on Affective Human Factors Design, Singapore, June 27-29, 2001: 317-324. Available at: http://www.webusability.ch/articles/CAHD2001.htm

Ferretti, F. (2009): The Credit Scoring Pandemic and the European Vaccine: Making Sense of EU Data Protection Legislation, 2009(1) Journal of Information, Law & Technology (JILT). Available at: http://go.warwick.ac.uk/jilt/2009_1/ferretti

Ferretti, F. (2015): Credit Bureaus Between Risk-Management, Creditworthiness Assessment and Prudential Supervision. EUI Department of Law Research Paper No. 2015/20. Available at: https://ssrn.com/abstract=2610142

Fogg, B. J. (2002): Persuasive technology: using computers to change what we think and do. Ubiquity 2002.

Gandy, Oscar H. Jr. (1993): The panoptic sort: A political economy of personal information. Boulder: Westview.

Gandy, Oscar H. Jr. (2009): Coming to terms with chance: Engaging rational discrimination and cumulative disadvantage. Coming to Terms with Chance: Engaging Rational Discrimination and Cumulative Disadvantage. 1-240.

Gandy, Oscar H. Jr. (2012): Statistical surveillance: Remote sensing in the digital age. In K. Ball, K. Haggerty and D. Lyon (eds), Routledge Handbook of Surveillance Studies. New York, Routledge, 2012.

Gandy, Oscar H. Jr. (2017): Neuroeconomics, Behavioral Economics and The Political Economy of Nudge. Available at: https://www.researchgate.net/publication/319942697_Neuroeconomics_Behavioral_Economics_and_The_Political_Economy_of_Nudge

Groupe Consultatif Actuariel Européen (2011): Use of age & disability as rating factors in insurance. Why are they used and what would be the implications of restricting their use? Position Paper, December 2011. Available at: http://actuary.eu/documents/GC_Age_Disability_Underwriting_Paper_051211.pdf

Hagiu, Andrei and Wright, Julian (2015): Multi-Sided Platforms. International Journal of Industrial Organization, Vol. 43, 2015. Available at: https://ssrn.com/abstract=2794582

Halpern, Sue (2017): How He Used Facebook to Win. The New York Review of Books, June 8, 2017. Available at: http://www.nybooks.com/articles/2017/06/08/how-trump-used-facebook-to-win/

Hanson, Jon D. and Douglas A. Kysar (1999): Taking Behavioralism Seriously: Some Evidence of Market Manipulation, 112 HARV. L. REV. 1420, 1564–65

Hartzog, Woodrow (2017): Privacy and the Dark Side of Control. The Institute of Art and Ideas, Sept 4, 2017. Available at: https://iainews.iai.tv/articles/privacy-the-dark-side-of-control-auid-882

Helberger, Natali (2016): Profiling and Targeting Consumers in the Internet of Things – A New Challenge for Consumer Law (February 6, 2016). Available: https://ssrn.com/abstract=2728717

Hoboken, Joris van (2016): From Collection to Use in Privacy Regulation? A Forward Looking Comparison of European and U.S. Frameworks for Personal Data Processing, In: Van Der Sloot, Broeders and Schrijvers (eds.), Exploring the Boundaries of Big Data, Netherlands Scientific Council for Government Policy, 2016, pp. 231-259.

Hoofnagle, Chris Jay (2013): How the Fair Credit Reporting Act Regulates Big Data (September 10, 2013). Future of Privacy Forum Workshop on Big Data and Privacy: Making Ends Meet, 2013. Available: https://ssrn.com/abstract=2432955

Hoofnagle, Chris Jay (2016): Federal Trade Commission Privacy Law and Policy - Chapter 6 Online Privacy (February 1, 2016). In: Chris Jay Hoofnagle, Federal Trade Commission Privacy Law and Policy (Cambridge University Press 2016); UC Berkeley Public Law Research Paper No. 2800276. Available: https://ssrn.com/abstract=2800276

IHS Markit (2017): The economic value of behavioural targeting in digital advertising. Analysis on behalf of IAB Europe and EDAA. Available at: https://www.iabeurope.eu/wp-content/uploads/2017/09/BehaviouralTargeting_FINAL.pdf

Kaptein, Maurits; Dean Eckles, and Janet Davis (2011): Envisioning Persuasion Profiles: Challenges for Public Policy and Ethical Practice. 9/10 Interactions 66-69, 66; Available at: https://www.semanticscholar.org/paper/Envisioning-persuasion-profiles-challenges-for-KapteinEckles/fe5f2029df491bdea2cf46697b2e4145c1e226f2/pdf

Kennedy, K., Mac Namee, B., Delany, S. J., O'Sullivan, M., & Watson, N. (2013): A window of opportunity: Assessing behavioural scoring. Expert Systems with Applications: An International Journal, 40(4), 1372-1380. Available at: http://arrow.dit.ie/cgi/viewcontent.cgi?article=1024&context=scschcomart

Kramer, Adam D. I.; Jamie E. Guillory; Jeffrey T. Hancock (2014): Experimental evidence of massive-scale emotional contagion through social networks. PNAS vol. 111 no. 24, 8788–8790. Available at: http://www.pnas.org/content/111/24/8788.full

Linoff, Gordon S. and Michael J. A. Berry (2004): Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management. Wiley Publishing.

Lyon, David (2003): Surveillance as social sorting: Computer codes and mobile bodies. In: Lyon, D. (Ed.): Surveillance as social sorting: Privacy, risk, and digital discrimination. Routledge, London, New York.

Lyon, David (2007): Surveillance Studies: An Overview. Cambridge: Polity Press.

Marwick, Alice E. (2013): Big Data, Data-Mining, and the Social Web. Talk for the New York Review of Books Event: Privacy, Power & the Internet, October 30, 2013. Available at: http://www.tiara.org/blog/wpcontent/uploads/2013/10/marwick_2013_datamining_talk.pdf

Mirani, Leo (2014): How Facebook and Google are taking over your online identity. Quartz, September 26, 2014. Available at: https://qz.com/271286/how-facebook-and-google-are-taking-over-your-online-identity/

Newcomer, Eric (2017): Uber Starts Charging What It Thinks You're Willing to Pay. Bloomberg, May 19, 2017. Available at: https://www.bloomberg.com/news/articles/2017-05-19/uber-s-future-may-rely-on-predicting-how-much-you-re-willing-to-pay

Ngai, E. W. T., Li Xiu, and D. C. K. Chau (2009): Review: Application of data mining techniques in customer relationship management: A literature review and classification. Expert Systems with Applications 36, 2 (March 2009), 2592-2602. http://dx.doi.org/10.1016/j.eswa.2008.02.021

Nissenbaum, Helen (2004): Privacy As Contextual Integrity. Washington Law Review. 79.

O'Neil, Cathy (2016): Weapons of Math Destruction

Ohm, Paul (2009): Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization (August 13, 2009). UCLA Law Review, Vol. 57, p. 1701, 2010; U of Colorado Law Legal Studies Research Paper No. 9-12. Available: https://ssrn.com/abstract=1450006

Packard, Vance (rev. ed. 1981): The Hidden Persuaders

Pariser, Eli (2011): The Filter Bubble: What the Internet Is Hiding from You. Penguin Press, New York, 2011

Pasquale, Frank (2016): Bittersweet Mysteries of Machine Learning (A Provocation). Available at: http://blogs.lse.ac.uk/mediapolicyproject/2016/02/05/bittersweet-mysteries-of-machine-learning-a-provocation/

Peppet, Scott R. (2010): Unraveling Privacy: The Personal Prospectus & the Threat of a Full Disclosure Future (August 7, 2010). Northwestern University Law Review, 2011 . Available: https://ssrn.com/abstract=1678634

Pollman, Elizabeth and Barry, Jordan M. (2017): Regulatory Entrepreneurship (March 3, 2016). 90 S. Cal. L. Rev. 383 (2017); Loyola Law School, Los Angeles Legal Studies Research Paper No. 2017-29. Available at: https://ssrn.com/abstract=2741987

Rhoen, Michiel (2016): Beyond consent: improving data protection through consumer protection law. Internet Policy Review, 5(1). Available at: http://policyreview.info/articles/analysis/beyond-consent-improving-data-protection-throughconsumer-protection-law

Schiff, Allison (2017): More Than Half Of Age Data In Mobile Exchanges Is Inaccurate. AdExchanger, January 11, 2017. Available at: https://adexchanger.com/data-exchanges/half-age-data-mobile-exchanges-inaccurate/

Singer, Natasha (2012): Mapping, and Sharing, the Consumer Genome. New York Times, June 16, 2012. Available at: http://www.nytimes.com/2012/06/17/technology/acxiom-the-quiet-giant-of-consumer-database-marketing.html

Singer, Natasha (2016): When Websites Won’t Take No for an Answer. New York Times, May 14, 2016. Available at: https://www.nytimes.com/2016/05/15/technology/personaltech/when-websites-wont-take-no-for-an-answer.html

Solove, Daniel J. (2004): The Digital Person: Technology and Privacy in the Information Age (October 1, 2004). Daniel J. Solove, The Digital Person: Technology and Privacy in the Information Age, NYU Press (2004); GWU Law School Public Law Research Paper 2017-5; GWU Legal Studies Research Paper 2017-5. Available: https://ssrn.com/abstract=2899131

Stanhope, Joe; Mary Pilecki; Fatemeh Khatibloo; Tina Moffett; Arleen Chien; Laura Glazer (2016): The Strategic Role Of Identity Resolution. Identity Is Context In The Age Of The Customer. Forrester, October 17, 2016.

Stoller, Matt (2017): Equifax Isn’t A Data Problem. It’s A Political Problem. Huffington Post, 09/13/2017. Available at: http://www.huffingtonpost.com/entry/equifax-credit-bureaus-reform_us_59b95627e4b0edff97187e7d

Swant, Marty (2017): Facebook Is Building Its Own Neuroscience Center to Study Marketing. Adweek, May 23, 2017. Available at: http://www.adweek.com/digital/facebook-is-building-its-own-neuroscience-center-to-study-marketing/

Taylor, L., Floridi, L., van der Sloot, B. eds. (2017): Group Privacy: new challenges of data technologies. Dordrecht: Springer.

Tene, Omer and Jules Polonetsky (2013): Big Data for All: Privacy and User Control in the Age of Analytics. 11 Nw. J. Tech. & Intell. Prop. 239 (2013). Available at: http://scholarlycommons.law.northwestern.edu/njtip/vol11/iss5/1

US Federal Trade Commission (2014): Data Brokers. A Call for Transparency and Accountability. May 2014. Available at: https://www.ftc.gov/system/files/documents/reports/data-brokers-call-transparency-accountability-report-federal-trade-commission-may-2014/140527databrokerreport.pdf

Valentino-DeVries, Jennifer; Jeremy Singer-Vine and Ashkan Soltani (2012): Websites Vary Prices, Deals Based on Users. Wall Street Journal, Dec. 23, 2012. Available at: http://on.wsj.com/Tj1W2V

Vinik, Danny (2014): Uber's Prices Surged in Sydney During the Hostage Crisis, and Everyone Is Furious. New Republic, December 15, 2014. Available at: https://newrepublic.com/article/120564/during-terrorist-attack-sydney-uber-imposing-surge-pricing

Yan, Jun; Ning Liu, Gang Wang, Wen Zhang, Yun Jiang, and Zheng Chen (2009): How much can behavioral targeting help online advertising?. In Proceedings of the 18th international conference on World wide web (WWW '09). ACM, New York, NY, USA, 261-270. Available at: http://dl.acm.org/citation.cfm?id=1526745

Zuboff, Shoshana (2015): Big Other: Surveillance Capitalism and the Prospects of an Information Civilization (April 4, 2015). Journal of Information Technology (2015) 30, 75–89. Available at: http://ssrn.com/abstract=2594754