Data Ethics – The New Competitive Advantage

Book · September 2016

CITATIONS
18

READS
539

2 authors, including:

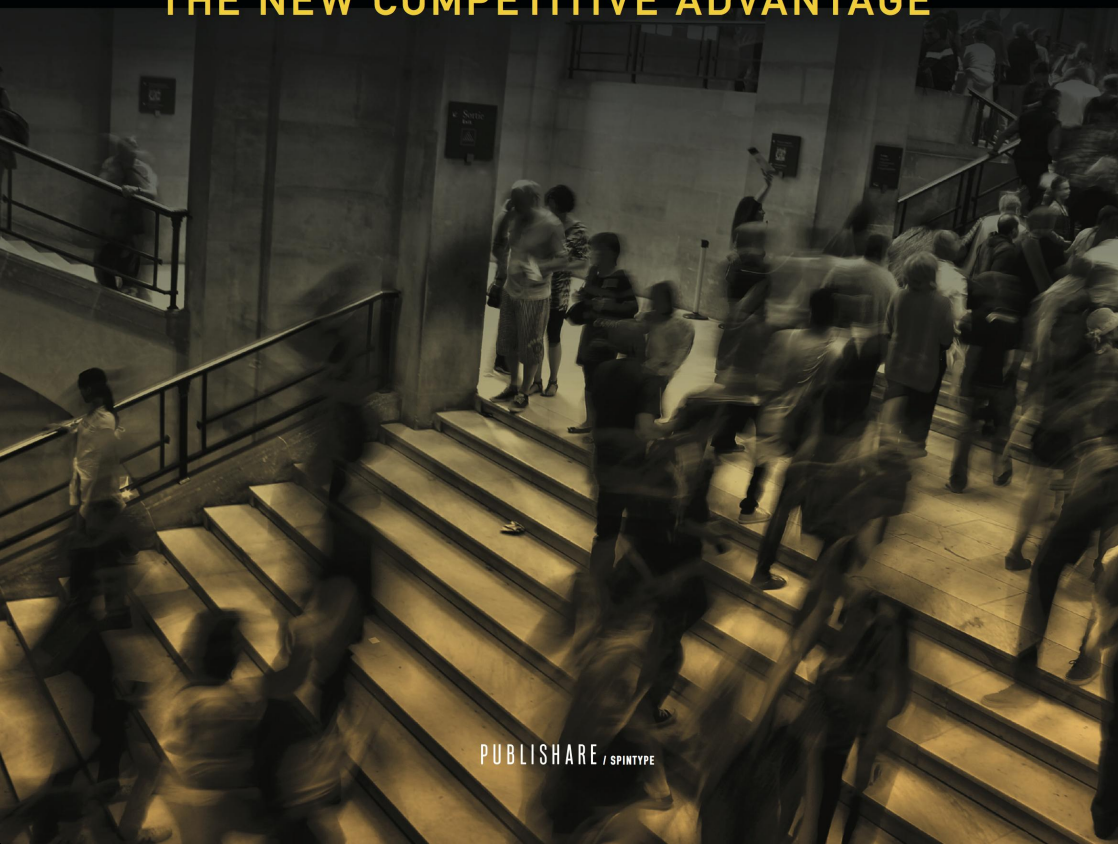
Gry Hasselbalch
DataEthics
9 PUBLICATIONS **67** CITATIONS

SEE PROFILE

Gry Hasselbalch & Pernille Tranberg

DATA ETHICS

THE NEW COMPETITIVE ADVANTAGE



PUBLISHARE / SPINTYPE

DATA ETHICS - THE NEW COMPETITIVE ADVANTAGE

1. edition, 2016

Authors: Gry Hasselbalch & Pernille Tranberg
Graphics: Publishare ApS / Spintype.com
Cover: Per-Ole Lind
Photos: Unsplash.com
Editor: Francesco Lapenta
English language revision: Katherine Kirby

Print: AKA PRINT A/S
ISBN print: 978-87-7192-017-8
ISBN pdf: 978-87-7192-018-5
ISBN epub: 978-87-7192-019-2

Supported by Internet Society

TABLE OF CONTENTS

Data ethics is potential, new market growth, a sustainable strategy and the foundation of creative, innovative business processes.

INTRODUCTION

THE DATA ETHICAL PARADIGM SHIFT

We are living in an era defined and shaped by data. Data makes the world go round. It is politics, it is culture, it is everyday life and it is business. Our data-flooded era is one of technological progress, with tides rising at a never seen before pace. Roles, rights and responsibilities are reorganised and new ethical questions posed. Data ethics must and will be a new compass to guide us.

Two decades ago, environmental reporting was something quite new, and many companies did not take being 'green' very seriously. There was growing concern among good-intentioned citizens, but many didn't know how to act on it. Today, those same worried individuals can sort their garbage, eat organic foods, take warm, solar-powered showers and drive electric cars. Companies also take the environment seriously. Not only because those with a direct effect on the environment are required to report to the authorities, but because green business practices are sound business practices.

Being eco-friendly has become an investor demand, a legal requirement, a thriving market and a clear competitive advantage. Data ethics will develop similarly – just much faster.

Data leaks, hacks, surveillance scandals and, especially, social media users' 'digital hangovers' (Chap. 1) have kick-started a movement. Individuals and consumers aren't simply concerned about a lack of control over their personal data (their privacy), they're starting to take action on it and react with protests, ad blockers and encrypted services (Chap. 3). In Europe, a new data protection regulatory framework which encourages the development of a privacy by default infrastructure has been implemented. Across the globe, we're seeing a data ethics paradigm shift take the shape of a social movement, a cultural shift and a technological and legal development that increasingly places humans at the centre.

Businesses are starting to feel this shift. Not as an 'either/or', either we use data or we don't, but rather they're gaining awareness about data from an ethical perspective, gradually moving away from an overbearing focus on big data (Chap. 5) and embracing sustainable data use. Visionary companies are already positioning themselves within this movement (Chap. 4) and investments in companies with data ethics are on the rise (Chap. 8). We're seeing an increasing number of businesses take the development of privacy technology as a direct point of departure (Chap. 6), along with the value of individual data control (Chap.12).

WHAT IS DATA ETHICS?

Ethical companies in today's big data era are doing more than just complying with data protection legislation. They also follow the spirit and vision of the legislation by listening closely to their customers. They're implementing credible and clear transparency policies for data management. They're only processing necessary data and developing privacy-aware corporate cultures and organisational structures. Some are developing products and services using Privacy by Design (Chap. 7).

A data-ethical company sustains ethical values relating to data, asking: Is this something I myself would accept as a consumer? Is this something I want my children to grow up with?

A company's degree of 'data ethics awareness' is not only crucial for survival in a market where consumers progressively set the bar, it's also necessary for society as a whole. It plays a similar role as a company's environmental conscience – essential for company survival, but also for the planet's welfare.

Yet there isn't a one-size-fits-all solution, perfect for every ethical dilemma. We're in an age of experimentation where laws, technology and, perhaps most importantly, our limits as individuals are tested and negotiated on a daily basis.

THE FOURTH INDUSTRIAL REVOLUTION

In the wake of today's rapid technological development, human and ethical dilemmas emerge (Chap. 11). Data is transforming society – some call it the Fourth Industrial Revolution. The first industrial revolution was based on water and steam, the next on electricity, and the third on information and digitalisation. In the fourth, the boundaries between the physical-biological and digital worlds are being eliminated – fuelled by data.

Data, personal data included, can have many positive uses and outcomes, but there are also many risks in a data-driven business process (Chap. 2). Gartner Inc. has predicted that by 2018, 50% of business ethics violations will occur due to improper use of big data.

GLOBAL STANDARDS FOR DATA ETHICS

Data is an asset, but it's also a risk. Today, the most prominent perils are data exhaust and unsustainable data practices, and a process to negotiate global standards, roles, rights and responsibilities to handle such risks has been initiated. This also means that tensions and clashes between laws and cultural values are amplified (Chap. 10).

Throughout history, societies have always somehow managed to mitigate man-made risks produced by different periods of industrialisation (e.g. pollution, atomic weapons and health hazards in food production) through new regulations, global standards, formal verification systems which consumers trust, and slow but steady cultural adaptation – including new levels of awareness, education, literacy and ethics. Industry has had to adapt to these requirements not only with targeted risk assessment and management, but by innovating and evolving in new ways. It will have to do the same in a data-saturated environment, with data ethics as a guide.

FAIRER MARKET CONDITIONS

There are several worrisome legislative circumstances worldwide that support indiscriminate mass surveillance of residents, back doors in technologies and greater secrecy shrouding intelligence services' monitoring activities. But there are also promising efforts which indicate a certain level of political understanding in relation to the privacy challenges inherent to the current digital infrastructure, as well as data's status as a new type of power. Although it's clear that many interests have had a say in the new EU data protection regulation (Chap. 9), it's still rather well thought out and attempts to look ahead to the technological evolution of the future. If enforced equally for both EU and non-EU companies and supported by anti-trust and consumer protection laws, there's a good chance that competition in the lucrative European market will be fairer than we have seen it the previous decade.

Regulation may point the way forward, but laws alone do not create fair market conditions or ethical business practices. Currently, companies can 'legally' use data in far more ways than what is in the individuals' best interest. Therefore, individuals must also take responsibility over their own data. It's a three-way hub of responsibility between regulators, individuals and businesses.

PRIVACY FOR THE ELITE

The societal repercussions of unregulated, ethics-free data practices are numerous, but the damage done to individual privacy is at its core. In a properly functioning democracy, those in power – government, industry and organisations -– are open and transparent about how they exercise their power. But one cannot expect transparency from individuals, as the more transparent people are, the more vulnerable they become (Chap. 13).

While laws, business practices, common international standards and cultural frameworks are being negotiated, privacy will be for the elite.

The highly educated, well-off, well-known and powerful will feel the need and will be able to pay for their privacy and control over their data. But as with the environment, a more formal framework for data ethics business practices will develop. A market for privacy tech and data ethics products will evolve, prices will go down and more people will gain access to them.

ABOUT THIS BOOK

This book is an analysis of trends through which we map a new field by looking at a few constructive solutions. This also means we address the forces at play in general, that is: the societal power structures, interests and relationships underpinning the field. It's fundamentally

important to us to make the invisible visible and, as such, provide the right tools to build something new: data-ethical services, businesses and products based on a paradigm shift in the way we approach digital data.

We hope to inspire companies large and small, as well as a wider audience of professionals who are not necessarily working in technology and data, but who wish to get a head start in the data ethics field. We have included more than 50 examples of practises that, in one way or another, are ethical when it comes to data. The examples were collected through interviews, credible media reports and website statements. We are not endorsing the companies, we do not compare their approaches, nor do we analyse all their practices. We are solely using them as case studies to provide the reader with inspiration for further exploration of the topic.

Most of the companies mentioned are still in a beta phase in the data ethics field, and not one has yet found the optimal solution. Every beginning takes time, just as it did with the products and companies that arose from the first inkling of environmental awareness.

Gry Hasselbalch & Pernille Tranberg, October 2016

In today's most common digital business model, consumers pay for 'free' products with their personal data.

CHAPTER 1

DIGITAL HANGOVERS

The year is 2006. *Time* magazine has awarded 'you' Person of the Year. *You* the active, productive web 2.0 user. *You* who use social media to share information, pictures and stories about yourself. *You* are hereby placed in the same category as Gandhi, Obama, Mark Zuckerberg and even the Earth: people and planets that throughout the years all have been named *Time*'s Person of the Year.

The 2006 award was recognition of online media's progress with the user at the centre. Social media, web 2.0 and active user centric services formed the most important trend in digital business development. Previously by invitation only, Facebook opened its social network up to everyone that year and Twitter launched as the first 'micro-blogging' site.

Traditional news media also jumped on the bandwagon. That same year, CNN became one of the first news media outlets to expand its services with *iReport*, inviting users to submit their own videos and photos from events around the world. Even savvy politicians found their very own channel in social media. In 2008 a relatively unknown man, at least from a global perspective, was elected president of the United States of America, partly based on a massive social media campaign and the use of data on the American electorate.

All of this happened because everyday people greeted social media with overwhelming enthusiasm. At first, it was trendsetters and young

people to use social media as part of their unique online identity. With photos, text and music they created online networks where they could coordinate social events with friends; they built a 'completely private' space sheltered from the prying eyes of concerned adults. It didn't take long before mum, dad, grandma and grandpa jumped on the web 2.0. train and began to 'poke' each other.

And what a party it was. The public media debate inspired somersaults of excitement for all the new opportunities: weblogs and moblogs, Youtube, Second Life, Myspace, Twitter, and something called Jaikuu. *You* were in the midst of sharing your life online. 'See my delicious menus, see my travels, see my baby. See me. Hear my opinion about shopping malls, lobsters and vitamins and politicians and skyscrapers and cars with three wheels! See, here I am at a party, so happy, loved by others...hey, that picture...can we delete that?'

OOPS, WE'RE ALL PUBLIC

It didn't take long before those same everyday people began to feel a bit of a digital web 2.0 hangover. Many had made a misplaced comment in the wrong context or posted pictures on social media that didn't quite fit their image. Parents began to check on their teens online and intervene in their social lives. We were misunderstood, some of us became enemies and some were even fired from our jobs.

Organisations, the media and politicians quickly shifted their focus towards the potential consequences of ordinary people suddenly becoming public figures with their lives freely available on social media. The original emphasis on Internet security evolved into a focus on responsible and ethical social media use. 'You are what you upload' declared one slogan. Then there was the danger of adult paedophiles, lurking in the dark corners of open social networking services. 'Never share your phone number', 'Don't talk to strangers online', children were taught. Another catchphrase reminded, 'You're the one who sets the limits', a mantra that, suddenly, web 2.0 users desperately needed to hear repeated; they were beginning to feel they had lost control.

Campaigns, safe chat rules, social media codes of conduct, guidelines and recommendations were all implemented. Everyone joined in, even the biggest social networking services themselves. Users were invited to adjust their 'privacy settings' and divide their networked friends into groups. One for colleagues. One for the family. One for friends. Public profiles were so last year. Facebook was seen as the most original trendsetter because they had private profiles, only accessible within the network, unlike its predecessor Myspace where your profile was viewable by anyone who stumbled upon it. The public debate began to slowly tune in to the more problematic aspects of web 2.0. We all had a web 2.0 hangover and we needed a cure. Social media services that didn't react fast enough lost the battle, or perhaps they just became necessary sacrifices in the first wave of public data ethics. Users lost confidence in the earlier social media sites and jumped on the bandwagon to ones which they thought they could trust (especially when their children were involved) and which offered them anything even slightly resembling control of their digital identities. We can all think of at least one open social media network that we used to have a profile on. (Did you remember to delete it?)

PERSONAL DATA BECOMES COMMERCIALLY VALUABLE

Parallel to the web 2.0 rush, two other trends were moving briskly along. The first trend was the foundation of Internet businesses after the first dot-com bubble. In the 1990s Ethan Zuckerman, the current head of the MIT Center for Civic Action, was working in one of the first Internet-based companies, Tripod.com. He offers insight into how an online business model based on targeted marketing came into being. Tripod.com had experimented with many different business models: subscriptions, shared user payments and even selling t-shirts. In the end, they landed on targeted marketing; meaning, as critics later pointed out, when something is free (or very cheap), you are the product. Under this business model, Tripod.com analysed users' personal web pages so they could target advertising to them. They chose

this route because it was the easiest one to sell to investors. As Zuckerman put it, the Internet was seen as "Christmas Eve for advertising and marketing people".[1]

The second trend was a society- and business-driven focus on data collection and storage, what is also referred to as big data. Knowledge, information and data have always been an important aspect of a company's business. But storing and analysing data before the age of the Internet required extensive resources. With the development of databases and analytical software, the cost of collecting and using data was significantly reduced. The Internet, web 2.0 and cloud computing, which made it cheaper to store data than delete it, created an additional foundation for what we now normally refer to as the data-driven business model, one founded on droves of data - big data.

BIG DATA RELIGION

Consumer tech giants such as Facebook, Google/Youtube, Amazon, Twitter, and Tencent have built their business models on the collection of data. Normally, they're described as social networking services and trading platforms, but they're also big data companies. Along with the more invisible data brokers that sell access to and trade data, they hold the world's largest troves of personal data with a growing range of applications. Data is at the core of their business models and processes, and these companies are assessed on the amounts of data they hold, their ability to put it to use and their capacity to innovate with it. The more data, the better. As Professor Viktor Mayer-Schönberger and the economist Kenneth Cukier posit in their book *Big Data* [2], the value of big data lies not only in the way we use data here and now, but also in the potential, future use of the volumes of data collected.The driving force for this type of big data business is the idea that large amounts of

1. Advertising is the Internet's Original Sin, The Atlantic, 2014.

2. Big Data: A Revolution That Will Transform How We Live, Work, and Think, Viktor Mayer-Schonberger, Kenneth Cukier, John Murray Publishers, 2013.

data equals great potential. It's an idea that can be translated into a type of big data religion or philosophy, as journalist Jacob Silverman argued in his criticism of the social media business model.[3] This corporate ideal is based on an almost metaphysical belief in raw data, where all data is seen as potentially useful and a potential road to success. In combination, the Internet, web 2.0 and big data business ideals evolved into a business model built on trackable aggregations of personal data. The resulting online infrastructure has a default setting which collects and stores data; it's a space where individuals are public and trackable by default.

The idea of big data has been a radically influential trend, not only in business development but also in science, governance, international development and surveillance. Methods based on the analysis of big data are being developed to manage natural disasters (by governments and humanitarian organisations), to trace the evolution of viruses across continents (by companies tracking the use of search terms), to follow electorates (by presidential candidates), and to predict individuals' future health situations (by insurance companies), potential criminal acts (by law enforcement), and the formation of romantic relationships (by social media scientists). Intelligence services are acquiring increasing access to archives of big data on citizens whom they want to keep an eye on.

Big data is just as great of a societal force of change as industrialisation was, and just as the industrialisation of societies brought about potential and growth, there are also many negative consequences.

The same can and will be said about the datafication of societies. We're slowly beginning to generate data through the things that surround us in our everyday lives. In 1984, science fiction author William Gibson described a virtual network, a cyberspace, he called it,

3. Terms of Service and the Price of Constant Connection, Harper Collins, 2015.

which citizens could connect, disconnect or be disconnected from.[4] In the 21st century it is increasingly difficult to fully log out of the online space. The concept of the Internet of Things (IoT) is used to describe the increasing number of objects in our environment which are connected to the Internet and which process data about us and our surroundings locally or in the cloud while we're at home or out and about. Smart cities, smart TVs, refrigerators, lamps, stereos, bracelets, watches and so on. Gartner Inc. has estimated that the Internet of Things in 2016 includes about 6 billion Internet connected objects and predicts that this number will increase to 20 billion in 2020.[5]

SURVEILLANCE REVELATIONS

Although big data already played a role in most institutional and industrial sectors in the years following the initial web 2.0 wave, it was a trend that few ordinary people cared about. More than anything else, the things that directly affected our personal lives and families were what created reactions. In other words, the immediate social challenges of being a everyday human, in public – when your boss followed your profile, when children exposed themselves online, when friends posted embarrassing pictures to a shared network. Users had their personalised web to worry about, while businesses had their targeted web to develop.

The digital hangovers which reflect an awareness of our data's secret life in a big data society are relatively new. They are the aftershocks of a series of events that illustrate specific risks associated with the storage of larger amounts of private data in the public global network. Most significantly, the episode with the most ramifications in the public sphere was whistle-blower Edward Snowden's 2013 revelations about the US National Security Agency's (NSA) big data surveillance

4. Neuromancer, The Berkley Publishing Group, 1984.

5. Gartner Says 6.4 Billion Connected "Things" Will Be in Use in 2016, Up 30 Percent From 2015, Gartner News Room, 2015.

methods. As early as 2005, with the book No Place to Hide, *Washington Post* reporter Robert O' Harrow had described, in detail, the dangers of the growing commercial collection of private citizens' data, combined with US intelligence services' increased focus on big data monitoring methods after the attacks on the World Trade Center in 2001. The first documents that Snowden revealed showed how the NSA collected hundreds of millions of text messages, email addresses, contact information and locations of citizens worldwide, every day. The documents also describe the so called PRISM programme under which, since 2007, the NSA collected data on foreign citizens via nine major US Internet companies, including the biggest social media companies. This disclosure about mass surveillance also proved to be a significant new angle on the framework for transferring data between Europe and the US and, as follows, the economic cooperation between the two.

It began to dawn on people that web 2.0 was not only big data, it was also big brother. In the months after Snowden's revelations, users of traditional privacy protection digital services mushroomed. Anonymous search engine DuckDuckGo's users rose by 50% in 2013, encryption tools of Silent Circle grew by 400% in weekly sales, while the encrypted cloud service Spider Oak's footfall increased 150%. The negative consequences of big data had become painfully obvious.

> **SpiderOak.** With SpiderOak, you can store data, collaborate with others and backup data. The service is based on the zero-knowledge principle. This means that SpiderOak knows nothing about the encrypted data, which is not decrypted until you use a password on your own computer. The customer therefore has full and genuine control over his or her own data. It's not just end-to-end encryption, which can leave behind information such as so-called metadata. It's zero knowledge, according to the company.[6]

6. Products with Principle, Spideroak.com, 2016.

There are as many good ways to use data as there are ways to create data exhaust and contamination.

CHAPTER 2

THE DATA DRIVEN BUSINESS MODEL

Google's search algorithm and Facebook's news algorithm are as guarded and coveted today as the recipe for Coca-Cola was in the 20th century. Such trade secrets are precious to companies in an online market which has become one of the present era's most financially lucrative spaces – especially for the fastest innovators and implementers, and not least those who understand how to scale globally. And for most large companies betting on the online market, data is the currency, means of payment, and foundation of their business models.

The first digital cash cow was the banner-ad. It was the first device which online news outlets used to generate payment for the content they published. With banner ads, sites could promise advertisers access to 'people North of London'; it was a way to reach a specific, targeted group of customers within a geographic area. Google, which did not produce significant income for the first seven years, began to capitalise on its search engine to then generate large parts of its revenue from banner ads. They could go even further and match those same 'people north of London' with their interests according to their search history, which Google stored and categorised. For years, Google and numerous other companies capitalised greatly on the search term advertising model (Google Adwords). Yet a new competitive model soon entered the playing field: an online social network which not only correlated demographics and interests but also people's real life identi-

ties and networks, precisely because users had to sign in with their real name in order to use the service. Although it didn't make a profit the first five years of its existence either, Facebook could go further than Google and connect the dots of 'men who drink red wine are owners of a caravan, heterosexual and single' for advertisers. Today, both Google and Facebook are among the most lucrative online companies on the planet.

The main tool to gather user information, cookies, has been refined over the years, and Google and Facebook have taken the lion's share of the advertising revenues based on cookies. While traditional news media are left to fight over the digital giants' leftovers, other, more quickly evolving companies are inventing new ways to harvest personal data to build detailed individual profiles. Yet at the same time, cookies are becoming an endangered technological dinosaur. They're losing steam as people are beginning to effectively block them both with ad and cookie blockers. More recent tracking methods are, for example, device fingerprinting, where you can precisely identify and track user behaviour through knowledge about the devices and applications they use, the size of their device's screen, time zone, fonts, etc. At the same time, information such as location and other relevant personal data is mined from mobile apps and wearables measuring one's health. Not to mention the upcoming data harvesting embedded in IoT (the Internet of Things), a business area in which the largest data companies have already taken root.

DATA AS PAYMENT

Many consumer tech and social media giants have built their business models on personal data. They may be search engines, social media, digital trading platforms, streaming services and health trackers. But they are, more than anything else, big data companies that generate profit on personal data. Although not necessarily trading data directly, their currency is similar to the more hidden (and often even richer) data brokers and data analytics firms. Data brokers such as Acxiom,

Datalogix and Experian trade individual data profiles while data analytics firms such as Palantir crunch data for the US government in the hunt for terrorists and large-scale fraudsters. In combination, these data giants have some of the largest archives of personal information on citizens all over the world.

Many of the consumer-directed companies have something in common: their services are either very cheap or 'free'. You simply pay an invisible price with your data, by now the web's preferred payment method. Modern-day customers have grown accustomed to not spending real money for digital products and services. They pay with data. Consequently, the digital companies of the future will find it even harder to profit unless they too offer their services for free.

The free model of payment even applies to businesses, which may opt to use the free version of Google Analytics. They pay, however, with customer data – in other words, their company's control over customer data.

GOOD DATA

Personal data is, at its essence, people, but data has also been defined as today's raw material, modern day gold or oil. Almost all major consulting firms have at one point used these terms to describe data's role in the current business economy. Many governments are betting that their countries' economies will evolve and grow based on data. Looking further down the road, this strategy is not as problematic as it sounds from a privacy advocate's point of view, because the most profitable types of data are not those connected to individuals. As McKinsey reported in '*The Internet of Things: Mapping The Value Beyond the Hype*' (2015)[7], the biggest growth potential lies in data which is not personal,

7. James Manyika, Michael Chui, Peter Bisson, Jonathan Woetzel, Richard Dobbs, Jacques Bughin, and Dan Aharon, McKinsey Global Institute, 2015.

such as that relating to weather, traffic and products, which can streamline production, logistics, distribution and service: "While consumer applications such as fitness monitors and self-driving cars attract the most attention and can create a significant value, our analysis shows that there is even greater potential value from IoT use in business-to-business applications. In many instances, such as in worksite applications (mining, oil and gas and construction), there is no direct impact for consumers."

McKinsey estimates that up to 70% of the value the Internet of Things is expected to generate in 2025 will come from B2B use of data. However, it will not be without the use of personal data and will not be entirely unproblematic, because the value in some cases is estimated to be even higher if personal data is combined with B2B data. Employers could, say, monitor their employees' blood pressure or blood sugar to keep them optimally 'maintained'. But in terms of B2B data, the biggest headline is the fact that data can optimise production and be used for 'predictive maintenance'. By way of a real-life example, Harley Davidson has a system which automatically adjusts to humidity and other conditions so that their motorbikes are coated with perfect enamel. On the topic of predictive maintenance, McKinsey describes how a company could prevent the collapse of a produced item due to damaged parts by monitoring the machines in real time and repairing parts before they break.

Vestas. The Danish wind energy company Vestas – the largest provider of wind turbines in the world – is a frontrunner when it comes to the use of big data. Before Vestas built its supercomputer with 15 years' worth of data on wind and weather, it could take up to 18 months to erect wind turbines which were optimally positioned in relation to wind and wind production. Today, Vestas uses an algorithm to create a statistical basis for decisions regarding the layout of wind turbines; the work is done with the click of the mouse. Drawing on data from 35,000 public stations which supply measurements

on over 150 parameters about every 6 hours, Vestas can produce accurate forecasts for long-term energy production at any point on the globe. Big data is also a part of their ongoing service. Sensors on individual wind turbines along with weather data are used to predict the wear and tear – and plan for the turbine's upkeep.[8]

Food Genius. This company delivers trends and data analyses to the food industry and is built entirely on big data. It retrieves data by scanning more than 87,000 menus from more than 350,000 American restaurants for a total of 50 million meals. This allows the company to analyse the diffusion and use of certain foods and menus as well as facts about individual ingredients, cooking methods and food types, such as organic or spicy dishes. The food industry uses the service to adjust production, develop and name new products, or change the menu.[9]

Enevo. Based in Finland, this company has developed algorithms which can foster more efficient waste collection in smart cities. Sensors inside the garbage container lid measure how full they are, so drivers don't use fuel in vain to empty half-full containers. By analysing the data collected from the sensors, Enevo can predict when containers are full and empty them accordingly. The stated objective is both environmental and economic, and so far the company's pilot tests in Helsinki and London show between 50% and 90% savings on driving, depending on the efficiency of the existing waste system.[10]

8. Datadreven vækst i Danmark, p. 16, IrisGroup, 2014.

9. Datadreven vækst i Danmark, p. 23, IrisGroup, 2014.

10. Harddisken, P1, DR, 2015.

iCow. Thousands of small cattle farmers in Kenya use the app iCow to optimise milk production. This program was invented by Su Kahumbu and provides the farmers with information about their cows' oestrous cycles, milking and market data.[11] In practice this means, for example, that the farmer receives a text message on the day the cow is the most fertile. The app collects the farmer's milk production and breeding data and sends him updated personalised advice and best practices via text messages, while simultaneously providing information on milk demand, veterinary data and market prices.

A growing number of organisations are developing methods to use big data for social or scientific purposes. They face similar ethical challenges as companies do, since they collect and store sensitive data, like that relating to health and location. Humanitarian organisations are using big data to trace the spread of a disease across a continent or to assess where to place aid centres. Big data is also used in food and medical research or to optimise the efficacy of hospitals.

Although the idea of using big data for humanitarian or scientific purposes is fundamentally different than the idea of profiting from its commercial use, the privacy risks are similar. In the report *Ebola: A Big Data Disaster*[12], Sean Martin McDonald explored the use of big data in humanitarian crises and the privacy implications for some of the world's most vulnerable citizens. As stated in the report, big data was even used to perform migration analysis and contact tracing without user consent during the 2015 Ebola outbreak in West Africa.

UN Global Pulse. Global Pulse is a big data innovation initiative from the United Nations. Its mission is to accelerate the discovery, advancement and scaled adoption of big data innovation for sustainable development and humanitarian

11. Udder genius: Fellows Friday with Su Kahumbu, 2012.

12. The Centre for Internet Society, 2016.

interventions. The initiative emphasises safety and responsibility, with a department dedicated to 'data privacy'. Global Pulse consists of a network of innovative labs and partners with experts from UN agencies, governments, academia, and the private sector conducting research in the field and developing new approaches and methods.

Just like some humanitarian organisations are exploring ways to use big data on individuals while respecting their privacy, a number of for-profit companies are also exploring ways to use big data without it pointing back to individuals. The focus here is effective anonymisation and secure storage.

Movvo is a Portuguese company that capitalises on location data. Movvo collects consumers' movements around shops with antennas installed in shopping areas. They then analyse the data and sell the results to retailers seeking to understand how consumers move around shops and cities. The company claims not to know who owns the mobiles they download data from, that they receive only a unique encrypted radio signal and that all data is anonymised. To gain trust from the public, Movvo obtained the Europrise privacy seal, which originally was established by the German national authorities and later privatised.

There is an embedded risk in collecting, storing and processing masses of personal data. Even when data is anonymised, the possibility of identifying individuals still technically exists if the data set is large enough. Some data which is not personal per se, can become so when correlated with other types of information. For example, if enough data is linked to Apple addresses that are in turn linked to individual consumer devices (Wi-Fi or Bluetooth identifiers), it is technically possible to use these identifiers to reveal who that individual is. For this reason, many companies are working on ways to use big data without compromising personal privacy. One of these is 'differential privacy', a

method which uses data hashing and noise injection to enable analytics on big data while keeping personal data private. In 2016, Apple stated that they plan to implement differential privacy for their services.

There are many good examples of using big data to support research and progress in developing countries, to streamline processes in factories and the like. Ultimately, challenges to privacy are often posed by the data collection and analysis' scope and, as follows, the proprietary context, how data is protected, and how it is stored and anonymised.

DATA AT RISK

The collection and storage of large quantities of personal data is in itself a risk to individual privacy. As such, data protection legislation in Europe prohibits collection without a specific purpose and requires user consent. However, the analysis and use of said data is, in particular, a challenge to individual rights. Algorithms are designed to make sense of data; algorithms are the foundation for data-driven services and the creation of profiles to personalise marketing and content, among other things. An online service can target advertising and content based on a person's preferences, previous patterns of consumption, and socio-economic background. These profiles can be so fine-tuned that they reveal intimate, private details, such as pregnancy or the likelihood a couple will get a divorce.

It goes without saying that there are ethical implications associated with the algorithmic analysis and use of data.

There are numerous examples of dilemmas to be found in the wake of algorithmic prediction, which may create opportunities or limitations for an individual. Algorithms and their design criteria have a direct impact on an individual's opportunities when, for instance, American prisons use predictive algorithms to calculate the probability that a

person will commit a crime again after his or her release from prison, based on various social and behavioural data on the detainees. The same goes for the use of algorithms by employers to match a potential job applicant's social media history with the company culture. For example, what are the selection criteria and which cultural and social indicators does it base its selection on? Are they biased? What are the privacy implications?

A risk analysis of the use of big data has been described by Professor Frank Pasquale in his book *The Black Box Society*.[13] He explains the effect of hidden algorithms acting on our digital data. These calculations can create or destroy one's reputation; they can determine a destiny. We do not, he argues, have insight into the motivations and intentions that lie behind them. We don't know how personal data is used, for what purpose and what the consequences will be for the individual.

Facebook's Newsfeed is an instance of the use of a black box algorithm according to a background paper produced for the government-funded Global Conference on CyberSpace (GCCS) in 2015.[14] Facebook's algorithm determines what you see on your wall and what you do not see. It's then adjusted by a team of researchers, who say they take thousands of factors into account. In 2016, it emerged that a small team of editors was actually injecting news topics into those trending among Facebook's users or into its 'blacklisted' topics. A Gizmodo blog claimed that this conduct was biased against conservative news items.[15] Facebook immediately rejected the claim, but it does not change the fact that what is presented as news to Facebook users is in no way a neutral presentation of accumulated user-generated content

13. The Black Box Society - The Secret Algorithms That Control Money and Information, Frank Pasquale, Harvard University Press, 2015

14. The Ethics of Algorithms: from radical content to self-driving cars, Centre for Internet and Human Rights, 2015.

15. Facebook's News Selection is in the Hands of Editors not Algorithms, Documents Show", The Guardian, 2016.

and interest in current events. Rather, it's a combination of algorithms and human editorial intervention, where the criteria and the processes behind such opaque editorial decisions are kept secret.

> **Target**. In 2011, the American discount retail chain Target developed an algorithm capable of finding customers who were about to become pregnant and even approximate their due date.[16] The chain could then directly market their baby and pregnancy products to these customers. The programme was so successful that the sale of products for pregnant women increased up to 30%. However, there was a problem: the creepiness factor. One day a furious father stormed into a Target location to complain that the store was sending pregnancy product adverts to his 16-year-old daughter. It was as if they wanted her to get pregnant. What he didn't know was that his daughter was already pregnant. The pregnancy algorithm knew before he did. A few years later, data on up to 70 million customers was stolen from Target, resulting in a major security breach and a sizeable economic loss for the chain.

It's not only risky for individuals to lose control over their data, data is also a risk for a company if not handled with due diligence and appropriate care.

> **Nets / IBM / Aller**. To many people's great astonishment, Danish publisher Aller's weekly tabloid, Se & Hør, managed, week after week, year after year, to reveal the buying habits and whereabouts of Danish celebrities, including the Danish Prime Minister Lars Løkke Rasmussen and the world famous actor Mads Mikkelsen. In April 2014 it was revealed that an IBM employee had been regularly texting data on Danish celebrities' credit card use to the tabloid's staffers. IBM was a subcontractor

16. How Companies Learn Your Secrets, New York Times, 2012.

of Nets Holding A/S, a payment and credit card service provider for all electronic money transfer methods in Denmark. As one of the small country's biggest privacy scandals, it gravely damaged the reputations of all three companies. In the end, it cost Aller a large amount of their readership and in 2015 the media organisation was reported to have lost more than 4 million euros.[17]

Mozilla. The owner of the popular Firefox browser promoted one of its best data analysts, Brendan Eich, to CEO in 2014. Six years earlier, Eich, who also co-founded Mozilla, supported an anti-gay marriage campaign with a contribution of $1,000. Back then the majority of Americans, including Obama, were against same-sex marriage. But by 2014 the national mood and opinion had changed, and when Eich's contribution came to light, the pressure against him became too heavy. He lasted just 11 more days as CEO, and Mozilla is now a cautionary tale of how vulnerable companies become if their employees do not also take care of their digital reputation.[18]

Samsung. In February 2015 a story about Samsung's Smart TV went viral.[19] 'Samsung spies on you', was the message. Journalists had been looking into the company's privacy policy and discovered that all conversations in the room where the TV was located were being recorded and processed by the company as part of a speech-to-text conversion service. The conversations were digitally delivered to a subcontractor to be processed into text form. Though it was unclear whether it was an opt-in or opt-out option, the viral discussion was incredibly unfavourable

17. Se og Hør sagen har kostet Aller 30 millioner, Berlingske Business, 2015.
18. How Mozilla Lost Its C.E.O., The New Yorker, 2014.
19. Your Samsung Smart TV is Spying on You, The Daily Beast, 2015.

for Samsung, which in the end responded with a 'we take privacy very seriously'.

DATA BROKERS IN A GREY AREA

There are plenty of examples of companies running into problems because they lacked control over their data and it's only a matter of time before more scandals emerge. Security breaches aren't only a threat caused by hackers from the outside or employees compromising data from within, harbouring direct criminal intent. Many companies also operate in legal and ethical grey areas in terms of what they can and cannot do with data. Insurance companies monitor customers and their data to reduce damage claims in court and assess their customers' health and behaviour to then adjust their insurance premiums. Airlines, car rental agencies, bookstores and many others establish prices based on knowledge about customers obtained by using cookies and other tracking tools. Many businesses both buy and supply data to the infamous multinational data brokers, companies that deal in personal information and which haunt the US in particular – or at least they're identified in the United States.[20] To illustrate this point, at the end of 2013 the director of the World Privacy Forum, Pam Dixon, disclosed that data brokers sell lists of chronically ill people, cancer patients, rape victims, alcoholics and the homeless to the pharmaceuticals industry.[21]

20. Data Brokers a Call for Transparency and Accountability, FTC, 2014.

21. What Information Do Data Brokers Have on Consumers, and How Do They Use It?, Testimony of Pam Dixon Executive Director, World Privacy Forum Before the Senate Committee on Commerce, Science, and Transportation, 2013.

A NEED FOR NEW BUSINESS MODELS

Throughout the history of the Internet, personal data (our location, identity and social conditions, consumption and behavioural patterns, desires, needs, interests) has been in the pipeline as part of a greater business movement focusing on direct marketing and personalised services. It's in this context that big data's potential has been somewhat misinterpreted as specifically associated with the storage and processing of personal data.

Erik Huizer[22] is CTO at the Dutch SURFnet. He was part of the web's infancy as both a developer and entrepreneur. His name is listed in the Internet Hall of Fame, which honours those who had a particular impact on the Internet's development. On the data-driven business model's development, Huizer had this to say:

"Nothing went wrong intentionally. People just started experimenting with it. People didn't mind giving away their data because they thought they were doing it in very specific contexts. To share information with people they knew. And then somebody else got the idea 'what if we combine this with other data?' This of course changed the whole privacy context without consent. They developed a business model without considering what this meant to privacy in general."

In the beginning, entrepreneurs and companies saw data as a monetary object, something that users needed to pay with in order to use their services. They didn't care about privacy. Later, there came a time in which companies said they did indeed care about user privacy, which, according to Huizer, was not very convincing as they were simultaneously collecting hordes of data on them. Today he cites a new, emerging trend in the Internet's technical and commercial development:

"Now we see the emergence of new companies that take privacy as a starting point. They structure their businesses from the beginning to acknowledge privacy and deal with privacy. Their business model is

22. Erik Huizer, November, 2015, personal interview.

based on an awareness of a backlash against the data monetising business model where users increasingly will flee towards their platforms. I've seen that movement over time."

The predominant digital business model, based on the raw harvesting and use of personal data mainly for the benefit of shareholders, is not only likely to have reached the lower limits of consumer confidence and corporate reputation. The model is also threatened by a growing number of users knowingly providing false data in the form of fake names, birth dates and spam email addresses in order to protect their privacy.

In the hunt for web traffic and downloads from new customers, a whole new industry has emerged (and is especially flourishing in Asia) which can supply anything that appears to be user activity. Some studies show that over 60% of all traffic never sees human eyes but is rather so-called bot traffic generated by computer programs, and that 90% of a company's marketing budget for online advertising is simply wasted.[23] With this knowledge, it is clear that we need to look for different business models in the digital world. Fortunately, there are more and more companies taking a few for a test drive.

23. The Alleged 7.5 billion Fraud in Online Advertising, Samuel Scott, blog 2015.

People are not just concerned about the surveillance capabilities of new technologies. They are also starting to act to actively avoid it.

CHAPTER 3

WHAT CUSTOMERS WANT

Many websites greet their customers and potential customers with personal messages, offers and prices. Some find it helpful, others intrusive. Personalisation is based on our digital footprint, but it feels particularly intrusive when someone holds and uses sensitive personal information about us without our knowledge. It feels like a betrayal of trust. Advertisements appear on your Facebook wall for things you don't remember ever having shared on Facebook – diapers, Alzheimer's, offers of assistance from a divorce lawyer or dating opportunities. Or what about the pair of boots or travel destination that continues to chase you around the web, even if you've already purchased them or have long since found an alternative holiday destination? And what about the price you paid for the rental car, hotel, flight or book. Are you sure you got the best price? Why did it rise the second time you came around? Perhaps others got it for less?

Most people would like to decide for themselves just who knows exactly what about them and when. At a flea market we bargain about the price, but online the playing field is uneven. The seller often knows more about the buyer thanks to intense data collection. In the era of big data, there's been a shift in the control over information about us. We have less oversight and less control of the data that forms our digital identity – personal information such as name and address, diseases,

needs, dreams, data on our family and network of friends, our motivations, patterns and habits. This lack of control is something consumers are beginning to feel directly and respond to. In an online environment, trust between a company and its customers is delicate, and thus a long term strategy must leave space to listen to consumers' concerns, observe their actions and react in good time.

GENERAL CONCERN FOR DIGITAL SURVEILLANCE

There's a movement going on among Internet users which comes to light by comparing studies, statistics and trends: they're beginning to demand control over their data. Several studies asking Internet users directly about the importance of privacy, data security and control suggest that they place these things high when ranking digital needs. Though a global trend expressed differently from region to region, it's particularly evident among consumers in the US and Europe, where the digitalisation of public services and use of digital media is high and where data leaks and surveillance scandals have been in the public eye.

Generally speaking, there has been a change in how the world's citizens perceive challenges and risks to their privacy.

While privacy violations traditionally have been linked with state-sponsored surveillance activities, many have also begun to worry about private companies' personal data collection.

A global CIGI-Ipsos survey[24] from 2014 showed that 74% of people from different countries in various continents were concerned that private companies monitor online activity, collect data and resell it. Much of this concern is associated with a lack of transparency in corporate data use and, consequently, consumers' lack of control over their personal data. Another survey which covered over 8,000 con-

24. CIGI-Ipsos Global Survey on Internet Security and Trust, November 24, 2014.

sumers in five countries (USA, Canada, UK, France, and India) from the Columbia Business School/Aimia[25] showed that 85% of people wanted to know more about what the collected data is used for, 86% wanted greater control over their data, and 80% would only provide their data to companies they believe they can trust.

On this topic in Latin America, Eduardo Bertoni[26], Director of the Center for Studies on Freedom of Expression at the Universidad de Palermo in Buenos Aires, stated: "There is a similar concern as in Europe regarding corporate surveillance. It varies from country to country. But people are starting to think that the main actor that affects their privacy is not the government, but the business sector. Most people see the private sector as a foreign power spying on them. This is also connected to an increasing mood of anti-imperialism. If they see a US company doing something in their country, true or not, they see it as a foreign state in their country."

In other regions, the trend is less clear. Such is the case in the Middle East, where basic access to online services often take priority over the right to privacy. Hanane Boujemi[27], Senior Manager of the Internet Governance Programme for the MENA Region at Hivos, says that many respond with a shrug to stories of commercial and governmental digital surveillance. "The interest in the Middle East in the concept of privacy is not as big as in the European region. They are used to surveillance. It is something that is lived on a daily basis for these people."

Concerns about commercial surveillance are strongest among consumers in Europe. Here, the vast majority of citizens accept that data collection is part of the digital business model and a prerequisite for gaining access to many digital products and services. In fact, 71% accept this condition according to a Eurobarometer survey from June

25. What is the Future of Data Sharing?, Mathew Quint and David Rogers, Columbia Business School, Aimia, 2015.

26. Eduardo Bertoni, November, 2015, personal interview.

27. Hanane Boujemi, November 2015, personal interview.

2015.[28] But at the same time, only 31% feel they have control over their data and a solid 67% of those surveyed are concerned about lack of control. The Eurobarometer survey, which included 28,000 Europeans, also showed that:

- 7 in 10 people are concerned that their data could be abused or that it will be used for purposes other than what it was collected for.
- Half of survey respondents said they partially read privacy policies, one third said they never read them, while only around one in five reads them thoroughly.
- 7 in 10 people also say that privacy policies are generally too long, and 4 in 10 people find them too difficult to understand.
- A large majority of Europeans expressed the belief that a company must always obtain explicit consent to use their data.

WHO DO INTERNET USERS TRUST?

Surveys in Europe and the US show that Internet users mostly trust regulated industries over non-regulated industries. Hospitals, banks and, partly, insurance companies are high on the trust scale. Search engines, social media and news media, on the other hand, are often the industries that Internet users trust least.[29]

TARGETED ADS AND PRICES

While some consumers appreciate targeted advertising that matches their style and interests, others don't care for it at all. One thing is certain; personalised content and offers are here to stay, and many com-

28. Special Eurobarometer 431 Data Protection, EU Commission, 2015.

29. State of Privacy Report, Symantec, 2015.

panies are trying to follow in the footsteps of Amazon and Netflix, experts in the delivery of what they call 'relevant' recommendations.

There's a great degree of variation in what consumers say when asked if they want targeted advertising, but often the answer depends on the way they are asked. According to Eurobarometer, 4 out of 10 are okay with the fact that companies use knowledge about their online behaviour to tailor advertisements and content. When respondents to a survey from the Danish Business Authority and The Danish Society of Engineers were asked similar questions, but with more specific information on the tracking processes that led to personalisation, they were much more sceptical. 'Is it a good use of cookies to give you personalised offers from the page you visit?' Only 24% answered yes. 'Is it a good use of cookies to give you personalised offers from other websites you visit (meaning that the advertisement follows you from other sites)?' Only 10% answered yes here.

A Norwegian study asked directly: 'Do you prefer targeted advertisements? (27%) or random advertising (73%)?'[30]

Evidence suggests that personalisation quickly gets to the point of feeling like manipulation, and consequently a company should use it with caution.

TEENS WANT PRIVACY

There has been a tendency to attribute concerns over privacy to the older generations' perception of the role of privacy in society. But young people actually place much more value on their online privacy than many adults think. They're frustrated by the lack of transparency in what data is harvested and why, and resentful of the lack of control they have when that data is used in targeting activities which are seen as invasive and irritating.[31]

---

30. Personal data in exchange for free services: an unhappy partnership?, Norwegian Data Protection Authority, 2016.

31. See e.g Youth State, survey on UK 16 - 24 year olds from Adjust Your Set.

Detailed studies on young people's use of digital media show that privacy is alive and kicking. Danah Boyd, an American researcher and founder of Data & Society in New York, came to this very conclusion. According to her, young people draw upon a wealth of complex strategies to maintain privacy on social media. They do actually want to keep some things to themselves, even while social and active online. Boyd has appropriately named the kind of privacy young people manage on social media 'social privacy'. In the minds of this demographic, privacy is connected to social context. For example, if an image from a social networking profile is taken out of the context it was posted in and used in a different context, many will see it as a violation of their privacy. Even if they have originally shared it in a place where everyone has access to it.[32]

Millennials are in fact quite aware of their privacy on social media. For them, privacy is not about closed boxes with locks and keys, but about being in control. They know that having a private life online means having control of the social context things are shared in. The great lengths they will go to in order to hide things from their parents online is proof of this. Studies also show that their views on privacy change according to their needs.

Once they enter the labour market, it's no longer just parents, friends and teachers. Suddenly, those who they must shield certain information from expands to include potential employers and others.

A majority of young Germans between 18 and 29 years old (54%) are against online policies that require that you use your own name when leaving comment, while 81% of the over-59 German demographic finds them quite okay.[33] In fact, there is a rather large opposition to sharing personal data with companies among young adults. Nine out of ten youths in the UK, for example, do not give away their

32. Why Youth (Heart) Social Network Sites: The Role of Networked Publics in Teenage Social Life, MIT Press, 2007.

33. Most Germans in favour of compulsory real names online, The Local De, 2016.

data no matter the benefits, according to a KMPG survey of 18 to 24 year olds.

DEMAND FOR DATA CONTROL

Young peoples' experiences navigating online identity, privacy and social networks provide an important insight into future market requirements and drivers of innovation.

It's not just the youth who want to be empowered online. Generally, consumers are demanding more and more control over their data and their digital identity.

First and foremost, consumers need to feel that a company is able to take care of their data. If so, they are more willing to share data with the company. This confidence in the security of data held by tech-based services is in many cases more important than good customer service or customer discounts. Consumers wish to be informed when their data is stolen or lost, though too few feel they are kept in the know. A majority of Europeans do not feel sufficiently briefed about what their data is actually used for when requested by a website. The facility to revoke permission – to delete data – is identified as the single largest factor in encouraging consumers to share more.[34]

A digital trends report from Microsoft Advertising[35] looked at the consumer motivations which drive online behaviour in different regions of the world. It shows a growing desire among consumers to control their digital identities. For example, 57% want to be able to choose how long information they share stays online and 80% are interested in services that 'manage their digital identity'. It is, as the report states, a shift from a consumer focus on privacy as a way to hide

34. Fair Trade?, Amazeone.com, Sarah Hooper, Paul Kennedy, 2016.

35. Microsoft Digital Trends 2015 - The evolution of digital consumer experiences, Microsoft Advertising.

digital footprints, to a focus on control of data and thus of one's digital identity.

CONSUMERS ARE BEGINNING TO ACT

In May 2014, the European Court of Justice passed down a judgment that is already playing a crucial role in the individual right to control personal data and privacy. With what is referred to as the right to be forgotten, Google and other search engines were asked to process requests to remove links to content and actually delete those deemed irrelevant, false or outdated. Google and a number of other tech and media companies immediately went out of their way to criticise the decision, saying it would have a negative impact on freedom of expression and provide criminals and public personages an opportunity to have content removed that could be in the public interest. However, contrary to the many warnings, it turned out that 95% of all requests to remove links came from ordinary people, not from criminals, politicians or public individuals. In countries like France, Germany and the Netherlands, 98% of all requests were based on privacy concerns.[36]

BLOCKING COOKIES AND USING VPN

The right to be forgotten is one thing. The right not to be monitored or tracked by companies is yet another. The current digital business model, tracking-by-default, means that individuals have to take action if they do not want to be monitored and receive personalised advertisements and prices based on their data. A rapidly growing number of users are starting to act. For example, millions of Internet users are blocking online adverts with adblockers, such as AdBlockPlus or AdblockFast and cookie-blockers such as Disconnect.me. Especially in the Western world, a large amount of people use blockers (the lowest per-

36. Google Accidentally Reveals Data on 'Right to be Forgotten' Requests, The Guardian, 2015.

cent in Ireland, the highest in Sweden and Germany), and this new blocking trend is considered a serious threat to the ad-based business model. Millennials are especially adept at using ad blockers. According to a US survey carried out in autumn 2016, two out of three between 18 and 24 years old used them.[37]

More and more are starting to use VPN services also. They encrypt web traffic, so it's safe to work on a free and open Wi-Fi-connection, and they allow users to hide or select the origin of their IP address. According to a GlobalWebIndex survey from 2015, one in four have used a VPN service. However, this is not necessarily because it protects privacy, but also because it can provide access to film streaming services worldwide.

FALSE DATA ON THE RISE

Another way to protect one's privacy is to provide false names and data online, to use a pseudonym or an alias. More and more people are doing this, and the younger you are, the more you do it. Companies like Facebook and Google have real name policies (Google abolished its real name policy in 2014), which means that their Terms of Service (TOS) requires users to provide authentic data. If you don't heed these conditions, the greatest risk you run is that your account will be closed. At least 20% of account holders (Facebook's own figures) use names other than their own. This could be anything from political activists and transsexuals to CEOs and people who just want to be left alone. False data is used not only on social media, but also in surveys, particularly when one has to fill in fields to get a report on a website, participate in prize drawings, and the like.

In the UK, up to 60% of users intentionally entered incorrect personal data about themselves, such as a false date of birth, email, name and address, according to a survey among consumers by the research

37. Millennials At The Gate, Anatomy Media, Fall 2016.

company Verve.[38] Eight out of ten users cited concerns about privacy as their main reason, but many also said that they want to stop companies from sending them targeted advertisements. In another study, which looked at Internet users in eight European nations (Symantec's The State of Privacy, 2015), one in three people lied online to protect their privacy.

OBFUSCATION

As part of the fake data trend, we also see attempts to drown true data in fake or 'dirty' data. In their book, *Obfuscation*[39], American professors Finn Brunton and Helen Nissenbaum describe a consumer revolution based on the conscious use of confusing, misleading and false information to prevent surveillance and data profiling. To address this, they created the browser add-on TrackMeNot. It drowns internet users' actual queries on engines such as Google, Bing or Yahoo with a long strip of ghost searches. The same with AdNauseam, which Nissenbaum is also behind. When you go to a website, everything on the site is automatically clicked, drowning the actual behaviour of the user in hundreds of clicks. According to a global Aimia-survey[40] of 8,000 consumers in October 2015, 67% have done something to protect their data – including providing companies with fake data.

FROM LACK OF KNOWLEDGE TO RESIGNATION

There is still a great lack of knowledge among consumers about what is really going on with their data. In a Harvard Business Review study[41] of consumers in five countries (USA, UK, China, Germany

38. Consumers are Dirtying Databases with False Details, Marketing Week, 2015.

39. Obfuscation: A User's Guide for Privacy and Protest, MIT Press, 2015.

40. How Business Can Gain Consumers' Trust Around Data, Forbes, 2015.

41. Customer Data: Designing for Transparency and Trust, Harvard Business Review, Timothy Morey, Theodore "Theo" Forbath, Allison Schoop, 2015.

and India) from 2014, only 25% knew that their digital footprints revealed their location, and even fewer were aware they also contain searches and Web browsing history. Symantec's State of Privacy 2015 survey stated that nearly seven in ten people don't know how to protect themselves against surveillance.

Now, one would think that more knowledge would lead to action. Not necessarily. The Tradeoff Fallacy[42], a survey from the University of Pennsylvania in June 2015, showed that nine in ten Americans do not think it's a fair deal to pay with their data for a digital service. It was assumed previously that many people gave their personal data to companies because they were unaware of what was happening with it. Yet this study shows that the opposite can happen; that those who know what is happening with their data are actually more likely to accept a discount in return for providing their data. Why? Because, concluded the authors, they are acting with resignation in relation to being in control. Resignation happens when a person believes the undesired result is inevitable and when they feel powerless to stop it. So rather than being empowered by the knowledge of their data transactions, some feel it is pointless to try to gain control of the situation. Though ultimately, more than half wished they had never lost control in the first place.

PAY FOR PRIVACY

Working at Carnegie Mellon University, Italian professor Alessandro Acquisti has made a career out of studying online consumer habits and their so-called 'privacy tradeoffs'. In a series of experiments, he looked at the value people attach to their privacy when presented with the choice to pay for its protection in different ways. His conclusion was that there's no evidence showing that consumers generally don't care about their privacy. The value they attribute to their privacy is complex and subject to a variety of factors, such as their personal motiva-

42. The Tradeoff Fallacy, Joseph Turow, Michael Hennessy, Nora Draper, 2015.

tions and the way choices are presented to them. In one study, for example, he investigated if consumers would pay for privacy.[43] Participants were asked to use a specially-designed search engine to buy a pack of batteries or sex toys with their credit cards. When the search results only listed the online shops, the subjects were not interested in the privacy policies. They simply bought only the cheapest products. But if the search results also showed comprehensible information about the differences in the online shops' privacy protection policies, the participants paid 5% more on average for products from those with the highest level of privacy.

In other studies[44]shoppers in a department store could choose between receiving an anonymous gift card with 10$ for purchases and a gift card with 12$ that tracked purchases. Here, those given advance notice about the better privacy protection their choice would imply if they chose the card with less money for purchases, would be five times more likely to take this card than others without this awareness.

There is no doubt that a company is better off protecting its customers' data and only using it for specific purposes rather than disclosing, sharing or selling it to third parties. It's a personal arms race and certainly the companies that collect data and use it in a lawful and ethical manner will be tomorrow's winners. Ad and cookie blocking, the use of VPNs, and fake data are clear threats to the tracking-by-default business model. Effective ad and cookie blockers are significantly on the rise, as they are easy to use and may have obvious economic benefits, particularly if you know how to fool a website into thinking that you are a first-time user. The use of fake data will also grow, as we see it used among millennials. There is a gap between what consumers want – openness and knowledge about the use of their data – and

43. The Effect of Online Privacy Information on Purchasing Behaviour: An Experimental Study, Janice Tsai, Serge Egelman, Lorie Cranor, Allessandro Aquisti, Weis, 2007.

44. What is privacy worth?, Alessandro Acquisti, Leslie John, George Loewenstein, The Journal of Legal Studies, Vol. 42, No. 2, The University Chicago Press, 2013.

what they see businesses doing. Long, incomprehensible privacy policies where users lose their right to control their data without fully understanding what it is they are accepting is an absolute no-go. As with the environment, consumers will realise that they must do something to gain control over their digital identity, causing a rise in demand for privacy-enhancing products and services.

Visionary companies are taking extra steps to safeguard privacy, to secure and protect data in order to build trust among their customers.