

Chapter 2

Remove Fear and Conquer Resistance

Businesses need to emphasize the societal value of Big Data

People Are Being ‘Milked’

Your personal data is valuable. Companies like Facebook and Google are making money from something that’s not really theirs—our data. According to some, we are being ‘milked’, like in the sci-fi movie The Matrix. In this 1999 movie, people’s energy was used to support a post-apocalyptic robot world. Lateral thinker Jaron Lanier sketches similar scenarios but suggests an alternative view. In his 2013 book Who Owns the Future?, he argues that we shouldn’t see computers as passive tools, because in doing so we fail to understand how computers and human beings interact. Moreover, he wants people to reclaim their own economic destiny by creating a society that values the work of all industries and not just those with the fastest computer networks.

Whether we like it or not, data is playing an increasingly prominent role in our lives. And for some time now, this has been happening not just in communications, but also in a whole host of different areas. Using data (networks) means, for example, that traffic lights can be synchronized to traffic levels, that we can turn our thermostats on or off remotely using our smartphones, that our cars warn the dealership when their sensors indicate something needs to be replaced, and that the Amazon website can

advise us on products we may like based on our customer profile. In an ideal world, data is a fantastic source of ease, comfort, luxury, and efficiency. However, this ideal world does not exist, and the disadvantages of data applications have become increasingly clear in recent years.

Drawbacks

Looking back, this has always been the case: all new technology offers solutions, but at the same time, it almost always creates problems. One of the clearest examples is nuclear fusion. Its discovery may lead to completely clean energy generation—at least insofar as we are capable of controlling the process. The dark side of this technology, such as total annihilation by nuclear bombs, requires far less control of the process and is, unfortunately, all too familiar. An even simpler example is the first time humans made stone tools. They came with tremendous possibilities, some immensely positive, with chopping wood as just one example—but also gave their owners the ability to inflict far greater injuries on other humans than would have previously been imagined. Technology may not only have adverse side effects, but it is also dependent on the motives of those using it with the potential to be exploited for both good and evil.

There are also two sides to the new information society. Social media offers a formidable tool for people to organize themselves. The bright side of the coin is that individuals can take a collective stand against powerful interests and even organize revolutions. The flip side, however, is that governments and companies also have a formidable tool to monitor people closely via the Internet and identify non-conformist elements. This is a problem not only because it can result in unacceptable forms of control, but also because many governments are not particularly transparent in this respect, to put it mildly. As such, we are seeing that

the new information society also has drawbacks that are becoming increasingly visible. There are three elements to these drawbacks, which we discuss below.

Privacy Breaches

Big Brother. The term—a reference to the famous book *1984* by George Orwell, which describes an all-knowing government that monitors its citizens in everything they do—seems unavoidable when discussing the future of the information society. The concept is not even particularly strange in an age in which the media informs us nearly every day that governments and companies are neglecting and invading our privacy. In 2013, Edward Snowden became the undisputed symbol of the fight against an increasingly pervasive surveillance state. His story is well known: as a whistleblower, he publicly disclosed that, since 2008, the US National Security Agency (NSA) had broken privacy rules or otherwise overstepped its authority thousands of times. This mainly concerned the unauthorized bugging and tapping of the data traffic of Americans or foreign nationals in the US. There was mass outrage, including from other world leaders—some of whose phones had themselves been bugged.

The information released about the NSA's activities increased the level of resistance people had to governments' appetites for data and the manner in which this hunger was being satisfied. To many, the revelations seemed to be a confirmation of a vague feeling of unease they had for some time.

These feelings are directed not only towards governments, but also towards 'data-rich' companies such as Google, Samsung, Apple and Facebook. There is a reason why these companies regularly clash with supervisory bodies on how they deal with personal data. One of the criticisms is how these companies combine

data, such as in the case of Google, which has access to information from various services, like YouTube, Google+ and Gmail. Combining information garnered from different applications is only allowed in Europe (and some other continents) within the boundaries known as ‘purpose limitation’. This means that when personal data is collected for a specific purpose, this data may not be used for any other purpose unless the person in question has given explicit authorization.

Painful incidents involving marketing also fuel peoples’ negative attitudes. A case that crops up regularly in discussions on how data is used by companies is the story of an American father who only found out his teenaged daughter was pregnant when the retailer Target sent her coupons for pregnancy products.¹ The supermarket had used its sophisticated algorithms to determine her status based on other purchases. Needless to say, this did not go down well. For retailers, having access to data that can be used to extrapolate such real-world information is extremely valuable in the battle against their competitors, because it allows them to target particular customer wants and needs which, in turn, breeds loyalty. Research shows that during life-changing events such as a pregnancy, people have a tendency to change their behavior and consumption patterns.² The first supermarket to recognize that these life-changing events are underway wins the jackpot and may have gained a customer that will remain loyal for many years.

The teenaged daughter had some explaining to do at home. Since then she has—probably reluctantly—become something of a symbol of the possible adverse consequences of companies analyzing (personal) data. The Target example makes it painfully clear that human behavior can be predicted with startling accuracy, much more easily than most people would think. Information about our lives is everywhere, with people performing their own forms of analyses. For example, you can pretty much deduce from status updates on Facebook the odds of a

relationship coming to an end.³ This is no surprise to mathematicians. Stephen Wolfram, a famous mathematician and physicist, who developed among other things a next-generation search engine (Wolfram Alpha), put it bluntly, “People are more predictable than particles.”⁴

Stories such as the Target one are often shared at the water cooler where everything is thrown into the same pot, with only one possible conclusion: businesses and governments are evil. They don’t care about my privacy; their only interest is in making money out of me (in the case of businesses) or exerting more control over me (for governments). And if businesses only use Big Data commercially in order to increase sales, resistance and resentment will probably only continue to grow.

Lack of Transparency

The resentment people feel is not only because of painful incidents such as the one described above, but is also fed by the lack of transparency on the part of businesses and governments. Because they are not being open about what they are doing with our data, they are not only breaking the law in some cases (whether or not they are aware of it), but they are also stoking resistance. This lack of transparency generates an enormous imbalance in society’s information flows: businesses and governments are getting to know us better and better, but we have to continue to guess how they operate. The question is for how long society will be willing to accept this lack of balance.

Here is an example: at the end of 2014, certain smart TVs made by the Korean company LG were collecting information and forwarding it to the company’s servers. This included information on which channels the television owner was watching, when the owner changed channels, and which programs were being stored on any media connected to the television by

a USB plug. A blogger discovered this by researching the Internet traffic of an LG television.⁵ He initiated his research after seeing a commercial aimed at potential advertisers in which LG was promoting the possibilities of targeted advertising based on the user data collected. On the consumer side, LG was much less transparent and, according to the blogger, the television even sent information when the consumer had switched off data sharing. The defense given by the company afterwards was not convincing. The lack of transparency about the data collection—the company had not clearly informed buyers of the television’s functionality—resulted in significant criticism of LG. For LG, it resulted in the firmware having to be adapted and a lot of reputational damage.

Strategically, it makes complete sense that television manufacturers should want to know our viewing habits. By using the resulting insights they can, at least in theory, provide us with a better service. But obtaining information secretly is not the answer. When companies are not transparent about collecting data from us and their use of such data, and do not offer their customers the choice of opting out of sharing data, they run the risk that their customers will turn against them. In recent years, that risk has been made evident by a determined group of privacy advocates who are out to name and shame organizations that (seem to) misuse our data.

Making Money from Personal Data

For many companies, data is the new gold. By using it cleverly, they can introduce new products and services, organize their activities more efficiently and offer their customers tailored products. Since personal data is worth so much to companies, can just anyone turn it into gold?

To answer this, we first have to clarify who owns it. In some cases—for example, your own personal data—this is completely

clear. In other cases, it is more complex: who, for example, owns the data containing the information that you clicked on a certain link on a certain website at a certain time? Do you own that? Or is it the property of the organization operating the website? To date, such discussions have only taken place in the background.

If we were to assume that companies such as Facebook and Google are using data that is not their property, we could simply ask them to pay us for it. That sounds nice—at least it seems reasonable from the user’s perspective! However, don’t expect this scenario to become main stream any time soon.

One of the reasons why this will never become reality is that we already receive a (hidden) reward. We are already compensated for making our personal data available and for the tracks we leave behind. For example, we receive discounts from loyalty programs and linked (reward) cards. Or what about free e-mail? When Google announced at the time of the introduction of Gmail that it would provide all users with a gigabyte of free storage, there was much disbelief and many even speculated that it was an April Fools’ joke. But it was serious. A decade later, it seems completely normal, and we are barely aware any more that this service is a reward for the data we allow Google to use.

There is a second aspect. If companies were to have to start paying us for (access to) our data, this would be an extremely complex process, with an allocation key being required, because not all data would be of the same value. Although this could be resolved, experience shows that the majority of people do not want more complexity but, on the contrary, want more ease of use.

Attempts are being made to introduce new models, but the current initiatives appear likely to appeal only to small niche markets. One example is the company Cayova (an abbreviation of ‘capture your value’) which started a social network focused on people selling themselves as advertising targets. The public at large is not (yet) embracing this idea. Another initiative is the

startup Datacoup, which pays for access to social media profiles and credit card transactions. The resulting insights are sold to businesses and in return, the user receives \$10 every month. It is also possible to earn more if users are willing to give up their privacy almost completely. Luth Research reads its users' computers and smartphones, gaining insight into, for example, their search terms, surfing behavior and social media profiles, and also wants users to answer questions about their purchasing behavior. These users can earn up to \$100 a month.⁶

We do not see a great future for the large-scale commodification of turning personal data into cash, because it is just too complex, but also because the vast majority of users basically accept the current situation. Therefore, for now, companies such as Facebook and Twitter have a truly enviable business model. They are making money out of something that is not theirs: our data. We accept this en masse, but paradoxically, it also feeds our negative feelings about how companies deal with Big Data.

The Wrong Approach

Without a doubt, privacy is an extremely important issue with regard to Big Data. Nevertheless, there is something funny about the way we talk about it. We seem to end up with almost completely polarized views, from the proponents of the advantages of Big Data on the one hand and the advocates of privacy on the other. Is this stopping us from getting to the heart of the issue, and finding the right approach to ensure our privacy?

We can illustrate this by using an analogy of an event in a completely different domain. In 2011, Anders Breivik detonated a bomb in a van in the center of Oslo and some hours later he shot and killed dozens of people on the island of Utøya. It emerged later that the bomb was made from fertilizer. The ammonium nitrate in the fertilizer can be released pretty easily and, together with an

explosive such as diesel oil and a detonator, it can cause a relatively effective explosion. In the months before the attack, Breivik had bought six tons of fertilizer without attracting any attention or suspicion. With his farm as cover, the purchase of such an amount appeared completely normal.

What can we do with this knowledge after such a horrendous event? Should we ban fertilizer? Introduce a maximum amount that can be purchased in any one transaction? Set up a strict control system to monitor sales of fertilizer? Implement tight security measures around the purchase of fertilizer? Forbid the online publication of instructions on how to make a bomb out of fertilizer?

It is not surprising that these questions have hardly been asked, but it is worth noting that none of these suggestions would have prevented the attack as there is no failsafe way to stop people gaining access to either products or information. We should point out that Congress in the US is demanding tighter legislation in this respect.⁷ Moreover, fertilizer is a very common product and imposing severe restrictions around its sale in the hope of preventing it from being used to make a bomb is like using a sledgehammer to crack a nut. We do not want to make the normal everyday use of fertilizer almost impossible because we are afraid that one lunatic will misuse it.

Misuse

However, this is exactly what has been happening in recent years regarding privacy in the context of a society in which Big Data is playing an ever-increasing part. New privacy risks are emerging in a society in which almost everything is becoming measurable. The strange thing is that this is one of the few social domains where (legal) measures are being taken to restrict the 'normal' use of data based on the fear that someone will exploit its potential abnormally. The legislation is focused on limiting

the collection of data rather than preventing its improper use. In terms of the terrorist Breivik and his fertilizer, this would be the equivalent of legislating to restrict the purchase of fertilizer but focusing little or not at all on the prevention of bomb building and detonation.

A secondary effect of this peculiar focus is that, in the debate about privacy, we actually hardly ever discuss privacy and instead focus mainly on information security. In itself, there is nothing wrong with focusing on information security—indeed, proper information security is essential in this day and age when everything is connected with everything else—but it is not the core of the challenge we face with regard to privacy. What we should really be discussing is how to ensure that the privacy rights of individuals are respected in a world in which ever more information is being collected about us. Instead, we are limiting ourselves to the question of how this data should be stored and secured.

When our focus in the privacy discussion is entirely on strict conditions for the storage of personal data, we may even be impairing the preconditions required to properly protect our privacy, because we are not seeing the other aspects of the issue. This is because having access to personal data is often not even necessary to breach someone's privacy. Again, an analogy will help clarify this. Say Pete goes to the same bar three times a week; he's probably welcomed as a friend by the proprietor, who probably exactly knows what brand of beer Pete likes to drink and maybe also that he has to be protected from himself at the end of the evening. So far, so good. No one sees any privacy problems in this scenario. This may change, however, the moment Pete takes his children into the same bar for a coffee. Pete may not want the bar owner to show how well he knows him, and he definitely doesn't want the bar owner to reminisce about how Pete fell off his bike last week because he drank too much. This would feel like an enormous invasion of privacy. Note that the bar owner has not stored any of Pete's personal data.

Again, we jump to the world of Big Data. In this world, data is often only valuable because a useful insight can be obtained by combining data from various sources. A chain of furniture stores can pick up and send Wi-Fi and Bluetooth signals from smartphones and, for instance, establish that a phone has been brought into the store for the third time this week and its owner has paused in front of the same couch each time. This could result in a discount being offered without the name of, or any other personal data about, the smartphone's owner being known to the store. The proprietor of the bar can do something similar by giving Pete, as a loyal customer, a beer on the house. No advanced technology is required to do that. However, there is a big difference: hopefully the bar owner has his own set of rules and values about how he deals with his customers—and his customers' privacy. Information processing systems do not have this trait as a built-in feature. What is required is that the designers of these systems—and/or the analysts that work with the data—are provided with ethical guidelines on the use of data; and that use must be supervised.

In the above examples, we discussed the collection of (large volumes of) data, which are known as 'implicit identifiers' (for example, signals picked up from smartphones). It appears that discussions about this kind of data are rarely focused on how it is used, but almost always about the possibility of turning this anonymous data back into personal data, such as through the use of data matching or similar techniques. This de-anonymization process is known as re-identification. But what exactly is personal data? When is a point of data traceable to a person? The challenge in this respect is that a single point of data cannot be traced back, but has to be combined with other data. Whether or not a point of data is traceable depends on the circumstances. And let's not even get into a discussion about what exactly traceable means. Because what if the bar owner doesn't know Pete's name and address. Does this mean his data is untraceable?

Often, these nuances are neglected, with many privacy advocates focusing on limiting the storage of anything that even smells like personal data. Although their intentions are good—fighting for the rights of the individual—this approach cannot hold up in a world that is increasingly data-driven. It could even be damaging, for example, by hampering large-scale research into various pathologies.

The reality is that organizations are often not too interested in personal data and, in many cases, only want aggregated and statistical data. To continue the analogy with the bar owner: he doesn't need to know Pete's last name or where he lives, as long as Pete feels at home in the bar and returns as often as possible. Although it is not often acknowledged that the discussion on privacy is too narrow, the fear of the unknown is accompanied by a reflexive urge to return to existing patterns, as demonstrated by a petition signed by one hundred European scientists, which states:

“Technically, it is easy to relate data collected over a long period of time to a unique individual. Economically, it may be true that the identification of individuals is not currently an industry priority. However, the potential for this re-identification is appealing and can therefore not be excluded from happening.”⁸

Prohibition

Fear can sometimes override more analytical thinking, which brings us back to the fertilizer. Based on this reasoning, limiting the collection of data translates as prohibiting the purchase of fertilizer because people such as Breivik exist. This just doesn't make sense. If we want to create a world in which we can capitalize on the massive potential advantages of Big Data in a controlled manner, we cannot try to prevent all potential future abuse by prohibiting normal everyday use.

Only when we focus on creating insights from personal data in a controllable manner, rather than (only) worrying about storage, can we make real advances towards protecting peoples' privacy. Only then can we have a discussion about the ethics of the usages we agree or disagree with and organize proper supervision of the analysis and application, not (only) the storage, of personal data. We discuss this ethical dimension further in Chap. 8.

Conflicting Behavior

Earlier in this chapter, we included some examples of how companies' and institutions' appetite for data has resulted in a number of unwanted incidents. It isn't hard to imagine that a growing number of businesses and governments will face such incidents in the coming years. It therefore seems likely that public resentment will increase, rather than decrease. Such incidents have mainly occurred in the 'new economy' of Internet, computer, and electronics companies. It is these companies that have often thought longest and hardest about this theme and push the limits.

Many companies that matured in the 'old economy'—banks, insurers, energy companies, etc.—have hardly even begun to think about whether they are handling data properly and whether they are sufficiently transparent for their customers. Only now are they starting to connect data 'silos' and realizing (although they are probably not realizing the extent of the issue) that, in doing so, they may be taking steps that will face public resistance.

A good example of this was the explosion of outrage⁹ that followed the announcement by the Dutch bank ING of its intention to make customers personalized offers based on their transaction histories. It was only a pilot project, but all of a sudden, all 17 million Dutch people had an opinion on how ING was handling their personal data, and a pretty strong one at that. The

consensus in populist discussions, such as those on Twitter, was that the bank needed to stop the project. Immediately. A number of national politicians even threatened to close their ING accounts. A few days later, the bank succumbed to public pressure and withdrew the plan (for now).

Banks are probably particularly sensitive to this kind of incident since the financial crisis—it is not for nothing that, in 2013, five years after the start of the credit crisis, banks and financial service providers were still the least trusted sector.¹⁰ Even more damning was a survey of 10,000 millennials that resulted in a list of the top 10 most hated brands. Four of those brands were banks.¹¹ Other sectors are also facing similar criticism. In the UK, hospitals sold medical data to insurance companies on a large scale, understandably causing a public outcry.¹²

There is a big difference in this regard between the US and Europe. Americans are more accustomed to the fact that companies use their data for various purposes, and the use of data is therefore a much less sensitive issue. Nevertheless, even in the US, companies may face significant reputational damage if their zeal for data collection goes too far. A good example is the protest that occurred when it became clear that OnStar, a subsidiary of General Motors, which collects GPS data from cars—including on behalf of insurers—had changed its conditions. According to the new ‘terms and conditions’—which are often routinely accepted without being read by users—data collection would continue even after the user’s account had been terminated. When this became common public knowledge, there was an outcry, with politicians also getting involved. Senator Chuck Schumer said it was “one of the most brazen invasions of privacy in recent memory.”¹³

This is only a snapshot of the many indications out there of significant resistance to the new forms of data consumption. It is not surprising, therefore, that discussions on talk shows, in bars and at the coffee machine at work about such (privacy) incidents

often become heated. How is it then possible that companies and governments who have so grossly abused our data have gotten away with it? Edward Snowden was (and is) a hero to many and has opened our eyes. The gist of most peoples' reactions to his revelations has been that this has to stop.

However, something strange is often going on with these discussions. Five minutes later, we are back to talking about football or that funny viral on YouTube and checking our social media as if nothing happened. We may even post a selfie, or tag our friends so that they know which café we're at. We may even post how many beers we've had...

The moral of the story is that Snowden has (so far) changed very little about our behavior. We are angry, but we can file that anger away as a memory very easily. When an even cooler app, smartphone or smartwatch comes along, we'll probably want that too. We get worked up for a second and then it's back to business. How can we explain this conflicting behavior? Let's list some factors.

Everyone Loves a Freebie

"There's no such thing as a free lunch." These are the oft-quoted words of the American economist and Nobel Prize winner Milton Friedman. Many products and services may seem free, but they never are—there are always hidden costs. The 'free' economy has always existed but, with the rise of the Internet, it has become much more prevalent and is contributing to our contradictory behavior.

"\$0.00 is the future of business," said Chris Anderson as far back as 2008 in *Wired*.¹⁴ He differentiated six business models that provide free products or services. One of those, the advertising model, has become dominant on the Internet over the last

10 years. The content, software, or service is free, but the user is exposed to the advertisements.

The advertising model fits the nature of the Internet perfectly. First of all, the number of people who see a certain ad and the click on it can be measured precisely. Advertisers therefore no longer have to pay for ads or commercials that no one looks at. But a second aspect is at least as important: the Internet makes tailored advertising possible. Because Internet companies can now learn about your likes and dislikes, they can tailor the adverts to your taste. The advertising model—invented by newspapers and magazines—has therefore been refined on the Internet over the last 10 years. The late Freddy Heineken, former chairman of the board of directors and CEO of the brewery Heineken International, allegedly used to complain that half his advertising budget was wasted but that the real problem was that he never knew which half. The Internet has at least eased that pain somewhat.¹⁵

For us as consumers, it's also good news. We don't have to hand over any money to use the best search engines, music services or handy apps. We can pay to use these services with our data. However, we need to realize one thing: if something is free, most likely we are the product.

In this regard, it is intriguing to look at the belief held by Google founders Larry Page and Sergey Brin when they were studying at Stanford University and how this has evolved over the years. While at college, they wrote, "We expect that advertising-funded search engines will be inherently biased towards the advertisers and away from the needs of the consumers."¹⁶ Compare this to the 2014 Google philosophy: "Focus on the user and all else will follow."¹⁷ This is in huge contrast to their vision as students, when they did not think it would be possible for the user to be central in the advertising model.

We Are Creatures of Habit that Follow the Herd

Humans are not always rational beings making well-considered choices. Often, we're creatures of habit, stuck in routines that are nearly impossible to change. This is part of the reason why people find it so difficult to break bad habits such as smoking, eating junk food and not exercising enough. We know that we have to change and we will change. But not today. Tomorrow.

Routines are essential to our lives due to the limits of our brain capacity and, because we are basically creatures of habit, we are comfortable following these routines. Scientists claim that our brains are continuously searching for ways to lighten the load and our habitual behavior is very helpful in this respect, because habits don't require much brainpower. The book *The Power of Habit: Why We Do What We Do in Life and Business* by Charles Duhigg describes how a woman turns her life around completely. She quits smoking, changes jobs and starts running marathons. Scans show that the patterns in her brain have also fundamentally changed. But basically, her brain has exchanged one autopilot setting for another.

We are not only creatures of habit, but also pack animals. This is the cause of a significant amount of both fraud and questionable business practices. The reasoning, whether conscious or not, is as follows: if everyone else is crossing a line to get a customer on board, it's okay for me to do it too. This was the case before the last economic crisis, when almost everyone in the financial world seemed to think there were hardly any risks involved in financial transactions anymore; and it was almost impossible to disagree as an individual. The habitual behavior in our immediate environment—that of colleagues, customers, etc.—greatly influences our own norms and values. People mostly learn their norms from the behavior of people they see as role models. This is aligned with the theory of cultural relativism,¹⁸ which tells us that there is no universal truth on which to base our ethics, but

that our interpretations are significantly influenced by culture and therefore our environment. We have almost all been the victims of herd behavior at one time or another. The Dutch brain research scientist, Victor Lamme and the late Harvard professor of psychology Daniel M. Wegner, even go so far as to say that free will does not exist.¹⁹ According to them, we mainly imitate other people and follow the instincts of our brain, while fooling ourselves into thinking that there are explanations for our behavior. In reality, however, according to Lamme, our thoughts follow our actions and not vice versa. One of the examples he uses is the cards you find in hotel bathrooms asking you to help save the environment by leaving only dirty towels on the floor, and towels that are still clean on the hook. The more clearly the cards show how many guests respond positively to these requests, the greater the willingness of others to cooperate.

We Are no Better Than the NSA

According to some, we are already living out the classic doomsday scenario of a surveillance state, with a multitude of surveillance cameras on the streets and disconcerting government tapping practices. However, the reality is that we are completely complicit in this. George Orwell's 1984 Big Brother scenario assumed an all-powerful government that had eyes and ears everywhere. The reality is that we are the biggest Big Brother. We continuously record our daily activities—and those of our friends—with our phones; we upload photos to Instagram or Flickr and videos to YouTube and Vine. On-board cameras in cars mean we can even 'enjoy' the accidents of our fellow road users (the latter being very popular in Russia in particular).

The real issue is not so much surveillance, but sousveillance. Surveillance is about an entity watching us from above ('sur' means 'from above' in French), while in sousveillance, the users transmit information upwards from below ('sous' means 'under'

in French). It became clear that sousveillance is a formidable tool during the search for the perpetrators of the 2013 Boston marathon bombings. A collaboration between users of the social networking site Reddit very rapidly resulted in a manhunt for a missing student who had, completely erroneously, been identified as the perpetrator of the attack.²⁰ These amateur detectives were not equipped to assess whether the ‘suspect’ was the right person and did not think what the consequences of their actions might be. Without any real evidence, their judgment had already been made. There is only one possible conclusion: we are turning ourselves into Big Brother.

We Are Addicted to Personalized Information

We expect Google to provide us with relevant and useful information that exactly matches our personal needs. In order to do this, Google needs to know more about us. The more it knows, the more relevant the information it can give us. From a music service such as Spotify, we expect our favorite playlist to be available on all our devices and maybe even to receive advance notice of a concert nearby that we may like to attend. For this to be possible, we have to share data with Spotify.

Personalized, tailored services seem to be the norm for every service provider on the Internet. We have become addicted, and our expectations get higher all the time. Ideally, we would like our search engine to know what we want before we do, our mail program to prevent embarrassing mistakes and our music service to have already ordered the tickets for a cool concert before anyone else even knows that our favorite artist is coming to town. Easy as pie. But these desires can only be fulfilled if we are willing to give all these companies access to our data.

Tension Between the Pros and Cons of Big Data

There is clearly tension between what we see as the pros and cons of Big Data. On the one hand, Big Data offers big advantages and we, as users, have massive expectations about how organizations should make those advantages available to us. On the other hand, there is a lot of resentment about the accompanying downsides and even the whiff of Big Data tends to meet resistance. It is human nature to resist change and this resistance is fed by the scope (and speed) of the changes—which are huge in this case and definitely disruptive. How can businesses and governments ensure they end up on the right side in this tug of war?

Offering Social Value

Remarkably, many (big) companies are still not making real choices in respect of the fundamental changes that Big Data is bringing their way and how they want to deal with it. Not choosing at all is the worst possible choice. They are holding onto what they know and are afraid of losing existing business should they decide on a (radical) change of course, into the unknown.

History has demonstrated many times what the risks are when companies are incapable of embracing new and disruptive concepts. Kodak missed out on the rise of digital photography, even though the company had the knowledge it needed to become successful in this area. The smartphone brought Nokia to its knees. WhatsApp made texting practically obsolete. Music streaming services turned the music industry upside down—or at least that part of it that was unable to keep up with a changing digital music landscape. It's possible that Bitcoin will do the same to banks. Bill Gates is quoted as saying: "We need banking but we don't need banks anymore."²¹ but more on this subject in Chap. 6. And what if self-driving cars actually became a fact of life for ordinary road users, not just a Google research project. This

would have an enormous impact on both business and industry. In such a world, would car insurance even be required?

These examples are closely connected to the new possibilities afforded by Big Data. The question remains, however: how can companies embrace Big Data without meeting resistance from the public and thereby shooting themselves in the foot?

In essence, the answer to that question is simple. At the start of a Big Data initiative, every organization needs to consider the value to the user. If only commercial gain is considered—how can we sell more diapers, advertisements, cars—or if the benefit of using the data becomes difficult to justify, sooner or later companies will end up having to deal with a serious incident. Therefore, companies need to be sure they are not only going to create financial value—make profits—but are also going to create social value, offering the user an additional benefit. Such companies will then have a future-proof strategy, tapping directly into the zeitgeist.

The prominent Harvard professor, Michael E. Porter, outlined this approach in an article in the *Harvard Business Review*²² with the title *Creating Shared Value*. In this concept, a company pursues not only economic profit but also value creation for both people and the planet. According to Porter, future businesses will adopt models under which they create social value that may also result in financial profit.

This is not a naive assumption, but a trend that is already visible from the early adopters. So far, the leading examples of such business models have had little to do with data. For example, Unilever is focusing on reducing child mortality due to poor hygiene in poverty-stricken areas in Asia²³ by providing special soap products and education. This gives the company a moral license to operate because, in addition to commercial value, it is also creating social value (and at the same time, demand for soap is increasing). The reaction of Apple CEO, Tim Cook, during

a shareholders' meeting in March 2014 was a clear case of connection with the new zeitgeist. A conservative investor criticized Apple's decision to invest in green energy projects, because they do not result in immediate financial profits. Cook was infuriated by such shortsightedness and had some advice for the investor: "If you want me to do things only for ROI (return on investment) reasons, you should get out of this stock."²⁴ Cook made it absolutely clear that he is targeting both social and financial profit.

Offering Social Value in Practice

This same combination should also be key when using data, but how does it work in practice? Here are some examples:

MasterCard wants to increase credit card security by only allowing payments when your smartphone is near your credit card.²⁵ The company is performing pilot studies using the geo-data of mobile phones (which, in many other cases, is leading to a lot of comment about privacy issues). Commercial profit: lower fraud costs. Social profit: safer payments.

Snapshot is a device provided by the American insurer Progressive that can be plugged into a car's diagnostic port (usually on the underside of the steering column). The device monitors driving behavior—speed, braking behavior, distance travelled and driving after dark—and the Snapshot data is then transmitted to the insurer. The advantage? The 'better' the driver, the lower the insurance premium. It seems a very smart way to (financially) motivate young drivers to drive more carefully. Commercial profit: a lower premium. Social profit: improved road safety.

If we are willing to share our lifestyle data with our health insurer, this may contribute to the early detection of health problems, or to scientific studies aimed at finding treatments and cures. Commercial profit: lower health care costs. Social profit: better health. We discuss this in more detail in Chap. 4.

A Public Backlash May Never Be Far Away

We acknowledge that it may not always be easy to transfer the combination of financial and social value from the drawing board to daily life. An incident involving the Dutch satellite navigation company TomTom provides a striking example. This company owns a wealth of (actual) information on the driving behavior of road users. Its users are happy because, based on their selected destination, TomTom immediately suggests an alternative route if it detects delays from the GPS measurements it is collecting in real time from thousands of TomTom users. Local authorities also use this information: it gives them clear insights into traffic bottlenecks and how to resolve them. This also benefits the driver.

In 2011 however, it emerged that TomTom was selling the aggregated data it collected about car speeds, via a specialist government agency to the police who said that it helped them to obtain insight into the most dangerous traffic situations. When it emerged that the information was also being used to determine where to place speed checks, there was nationwide uproar.²⁶ People were furious that TomTom had allowed this to happen, even though the data was anonymous and aggregated and no private information was involved. TomTom might believe it was providing a social benefit by contributing to greater road safety. However, the public thought otherwise. TomTom immediately stopped providing the data, apologized and changed its policy.

The commotion surrounding the use of payment data by ING mentioned earlier also shows how careful companies have to be. The bank received a tsunami of negative criticism after a test was announced that would enable the bank to offer personalized promotions based on customers' spending habits. ING is a symbol for the struggle that many other banks—in both Europe and the US—are going through. We should also point out, however, that a small number of opinion columns and weblogs internationally

praised ING, because the test was exploring ways the bank could create actual added value for its users. ING's clumsy and confused communication about its plans probably played a significant role in the resultant national outcry.

During a second attempt to explain its plans to the public at large, the bank emphasized how it was trying to understand customer behavior better, based on payment data, with the potential advantage of improved customer service. Was this second attempt a real clarification of its plans or more of a rebranding exercise? There's no way to know for sure. According to ING, it was never its intention to sell customer data to third parties—which was the issue that met so much resistance and continued to provoke so much online discussion long after the bank had backed down. The bank also announced that the plan would have been discussed extensively with customers and other stakeholders before any decision was made about its implementation.

The entire affair was a major public relations disaster for ING. It has left its customers with a bad aftertaste that will not fade overnight, even though ING claimed to be acting solely in customers' best interests. This case shows how vital it is for organizations to clearly communicate the intention behind their Big Data projects.

A Devilish Dilemma

It is clear that the wealth of insights into our actions that is available to businesses is becoming an increasingly important aspect in competition. Eric Schmidt, the former CEO of Google, put it strikingly in an interview in 2010 with *Der Spiegel* in Germany: he sees a future in which machines and technology play an ever-increasing role and he wants Google to know as much as possible about us, simply in order to improve our search results. "We won't need you to type at all. We know where you are. We know where you've been. We can more or less know what you're thinking about. [...] Google policy is to get right up to the creepy

line ... but not cross it".²⁷ Many Big Data plans face a major dilemma. Companies need more and more data in order to create (social) value to provide the services we require. The more we are willing to share that data, the more they are capable of meeting our needs and thereby creating social value. But many people are strongly against sharing more data. They associate Big Data with Big Brother scenarios, and are worried that companies only want to make as much money as possible from our personal data and that governments don't care about our privacy. We can't blame them, as these feelings have been fed in recent years by numerous scandals. Time and time again, it has been shown that it is essential for businesses and governments to make it clear to the public why they are sharing their data, so they can decide for themselves if this is a good thing. Those who can't do this are unlikely to be successful in the long run and may as well shelve their plans before they even start.



<http://www.springer.com/978-94-6239-182-6>

We are Big Data

The Future of the Information Society

Klous, S.; Wielaard, N.

2016, XIX, 199 p., Hardcover

ISBN: 978-94-6239-182-6

A product of Atlantis Press