Beuth University of Applied Sciences Berlin

Sprachenpreis 2016

Our Digitized Future:
Social Networking or Total Monitoring?

Written by

Friederike Buchner

Print and Media Technologies (BA)

Berlin,

September 16, 2016

Introduction: Data Collection Through Internet Services

Internet Corporations like Google or Facebook collect massive amounts of data. While they claim to do this to improve their services, centralizing personal information comes with inherent dangers. This is especially true because users do not seem to care. We will now proceed to show how and what data is collected. In this essay I want to take the reader through common practices of Internet giants like Google and Facebook. I argue that there is the social tendency to surrender the right to privacy in the face of assumably free internet services. This leads to potentially dangerous consequences. I will show the price we pay for using their services and give some ideas of how this may develop in the future.

Information serves as a currency for allegedly free online services (Kurz and Rieger 2011, p. 15). Services provided by Google like the search engine and Google Docs as well as Facebook's social network are indeed not free as one might expect. Every search request via Google's search engine is at the same time an answer (ibid., p. 65). We practically tell Google all about ourselves, our interests, our relationships, our needs. The same thing happens on Facebook. By evaluating our relationship status, answering friend requests and posting status updates we share our information with a much broader audience than we might be aware of. That means that a small action like sharing comes with the small price of giving away one's private data. One important difference to currency, however, is that the information in fact never flows. Internet companies tend to keep it for themselves.

Data never gets deleted and amasses quickly. Indeed, we can make entries disappear from our friends' timelines by pressing the "hide" button. But Facebook was never designed to make information go away. As Kurz and Rieger explain, user profiles are so enmeshed with each other that the question of deleting single elements, like for example that photo from five years ago that we might feel uncomfortable with now, would mean to delete the whole bundle of comments, likes and links to other users' profiles. To delete only that single photo would mean for Facebook to establish mechanisms, or personnel, whose only concern is to filter through such cases and make deliberate decisions of what stays and what goes away (ibid., p. 78). That would not only be very time consuming but also expensive for the company itself - especially since there is no interest whatsoever in letting the user monitor what information about him or herself should be available. With

Google it is a little different. More recent reports say that lately, Google allows users to selectively delete history entries (Sokolov 2016). But that is not for a postulated “Right to Forget” but for the benefit of a more accurate personalization. From this it seems that after information has been paid, it never gets quite repaid. It is safe to say that massive amounts of data get collected, but why does anyone care about who we date or what we shop?

The Creation of Economic Value Through Information

The big online companies have developed powerful techniques to extract meaning from the data. Through a practice called “scoring”, algorithms can quantify the information we leave behind about ourselves through credit card records, blog posts, mobility data, employment, etc. Then, they calculate an image of our personality based on a variety of parameters (Kurz and Rieger 2011, p. 59). For example, from my personal internet usage, Google knows that I am a white, female, twenty-something-year-old student, that I live in Berlin and go to the gym regularly, that I am interested in sustainable living, local and international politics and that I think the rents are too high. And since there is a bunch of other people fitting the same profile, algorithms can monitor our behavior, collect information of what the one likes and offer these as suggestions to the other. Through simple methods like the question “How helpful was this comment?” machines can learn to re-evaluate the suggestions they made for the user and appropriate the actual intelligence of people (ibid., p. 66). From there on they make predictions of what we will want and do in the future. These insights are being used for personalized ads, personalized search results and personalized product placement. Then the question remains: how do the companies benefit from this?

Having painted a picture of the user, internet companies can derive their revenue from two main sources: advertising and stock markets. A company’s value is determined through the number of new users and the amount of contacts they generate (ibid., p. 49). The complete data set of an active user of a social network with all the information and contacts he or she generated, can be sold for more than 20 EUR each (ibid., p. 44) on the company’s end. And the more ads she clicked, the longer she stays on a sponsored site and the more likely she is to buy a sponsored product, the more value her profile will gain. Advertising companies pay 1 EUR or more per click! That way, Facebook has

earned more than 3 Billion USD in 2011 through ads, which was approximately 82 % of its income (Kleinz 2012). But both Google and Facebook also directly monetize users profiles on the stock market. By evaluating the sheer mass of information, Google is capable of predicting stock exchange movements and recognizes which areas are going to be of investment interest (Kurz and Rieger 2011, p. 95). Of course, any operation of that scale will have vast societal impacts.

The Impact on Individuals and Society

The mass of information out there in the net is whittled town to an excerpt the user agrees on. Personalization does not only mean ads and product placement according to your presumed interests, it also means that when researching a term on Google's search engine, you are likely to get the search results Google *thinks* are best suited for you. Eli Pariser has dubbed this phenomenon "The Filter Bubble" (Pariser 2011). It means that search results vary according to the profile that Google designed based on your personal data. When searching for "Egypt" you may either be presented articles about the political crisis in Egypt or tourist information. Likewise you might not even see anything about the other topic respectively. In a lecture given to the London School of Economics and Political Science Pariser claims that "there is no standard Google anymore" (LSE 2011). Search results are ranked according to "relevance", not importance or verifiability.

Our search results only reflect our presumed interests but are far away from giving an objective picture of reality (Pariser 2011). As the "Mere Exposure Effect" states, people feel increasingly comfortable with things after only being exposed to it (Zajonc 1968). As in the case of social media, people tend to click more, view more ads and stay on sites longer when they see things that they are familiar with. As we have seen, this is just what social media companies earn money with. By controlling what we see, internet companies can reinforce the feeling of familiarity.

At the same time, people are not aware of what they share about themselves and with whom. Most of the internet users only have a very vague idea about how they are being monitored. Nor do they really understand what consequences that may have on society. With the net as an almost inexhaustible source of knowledge and information at hand, most people, particularly young people, use it as the number one source for educating themselves on current topics. When doing that, people automatically assume

that the net is neutral. But it isn't. As Pariser postulates, "there is no such thing as pure and neutral code" (LSE 2011). The ranking lists that Google produces are by their very nature already biased. And that leads to a kind of paternalism. We surrender control over our options. When in earlier times the producers of TV and radio shows were the gate keepers to information, it is now algorithms. The difference is that they work without ethics, without values. According to Pariser, the Filter Bubble was created around us through internet providers, but we, as users, have no influence whatsoever about what gets inside and we have no idea about what stays outside (ibid.).

It is a popular trope to highlight the chances that come with the influx of information. The "Big Data Dream" is being dreamed by law enforcement officials and disease prevention researchers alike. The hope is that through the accumulation of data it should be possible to determine statistical assertions about standard behavior and potential outliers. One example is the emerging trend of "The Quantified Self" where people track their own body functions and share their data with a big community. Deviations from an assumed healthy heart rate or blood sugar level are to stand out early, which provides a higher chance for an early diagnose of potential diseases (Swan 2013, p. 91).

Such crowd-sourced concepts could indeed lead to improvements in medicine and diagnostic tools, like the fact that this can fundamentally change our ideas of what being "normal" means and how able-bodiedness is perceived in society. If a lot of people participated in building up new norms through bringing in their own body measurements, our beauty standards might shift away from the underweight fitness model to the body size that the majority of people have. But bear in mind that these services for health evaluation are proprietary systems. That means, that the user pays for the service and a company provides it. There is no way for the community itself to access the data. Like in the case of "Quantified Self", there were services for health profiling available from 79 USD (ibid., p. 88). When the study was published in 2013 however, it claims that there was no open source project with a public database available, although it would have been "helpful" (ibid., p. 90).

Public access to such databases could well even be a necessity. As Boyd and Crawford 2012 argue, in the past it was always the most strenuous part of social studies to *collect* data. As a graduated anthropology student myself, I know that conducting research on social dynamics meant going out into the field, spending extensive periods of time getting familiar with the people and surroundings and only then being able to extract

data. That involved a huge effort to make sure not to “taint” the data through your personal interference. Today, a lot of data is already collected through social media. The problem is accessing and analyzing it. Of course, Boyd and Crawford point out that data from Twitter and Co. is not at all representing a comprehensive image of society (Boyd and Crawford 2012, p. 669). But however flawed, a sociologist working with Google or Facebook is able to conduct research on it while researchers from universities and other independent institutes are not. The possibility to work with those inexhaustible sources is indeed becoming a privilege for those with money or the right job. Furthermore, the broader scholarly community has no means to re-evaluate the claims made by those few (ibid., p. 673). But even if researchers got access to the data, they are unlikely to ask questions that might compromise the public image of the company that owns it if they fear repercussions (ibid., p. 674). To put it bluntly, long standing ideals of scientific transparency, objectivity and verifiability get seriously challenged by the current developments in social media.These problems exist today and they could get disproportionately worse in the future.

Considerations on Future Developments

Now, we have come a long way from how Google and Facebook mine data from user submissions, how they monetize it and how that affects us personally and as a society. But that is the development of data privacy in the now. I also want to have a look at how this could develop further in the future.

Mark Zuckerberg, founder and CEO of Facebook, has given a talk at the F8 Facebook Developer Conference and presented an overview about the company’s “10 year road map” (Cocktail 2016). To the critical eye and ear, our future looks scary and fascinating in the light of the visions that Zuckerberg provides. When describing artificial intelligence, he pictures software that is capable of watching and understanding photos and videos and being able to describe what is in it. This might be handy when it is about designing a barrier-free access to the net for the visually impaired. But it sounds scary when it comes back to the question of privacy. He himself points out that now, they (i.e. Facebook’s News Feed algorithms) are only able to use information such as who is posting the content and who is showing interest in it. They are not able to use the content itself. In the future though, he envisions, they will be able to read articles and understand

what they are about. All of that is supposed to serve the sole purpose of “showing you more interesting content from across the community”. Personalization is the keyword that he might be thinking about. Monitoring is the word that comes to my mind. When Facebook goes through with that we might face a potentially dangerous new step towards government surveillance. While it was meta data that we were concerned about since Edward Snowden’s revelations about the NSA scandal, it is now actual content that can be saved and monitored. And that will be data we have provided ourselves willingly, but not consenting.

There is yet another precedent being fought. It is between Apple’s CEO Tim Cook and the privacy legislation of the US about whether or not Apple can be obligated to give up its own encryption to surrender user data to law enforcement (Lichtblau and Benner 2016). It is no exaggeration to consider this one of the major dealings with this question and an important precedent. Personally, I do not like the idea of internet corporations being the guards of my information, but I like the idea of the government sifting through my personal conversations even less. Bear in mind that what has once been saved on the internet can be copied and reposted for an infinitive number of times. You can never assume that something has left the internet. In a way, our information is now permanent and that is definitely something to worry about. We trust the net with our daily conversations as well as our holiday memories as well as our frustration about the job. These could be recognized automatically and stored away in the server rooms of private companies. While the privacy of (analog) correspondence, posts and telecommunications is a highly respected value and achievement historically, that is an entirely different thing online. Economically and judicially, we are facing a future where personal information is not so much personal anymore. These developments seem to be supported by our political leaders, too.

In the Publishers’ Summit of the German Union of Magazine Publishers chancellor Angela Merkel gave a speech about digitization in the publishing industry (Merkel 2015). She emphasizes the importance of preserving the possibility to commercially process data rather than restricting it through data security legislation. She claims that the business model of creating value through Big Data Mining and data management deserves protection from too strict restrictions since “data is the raw material of the 21st century” (Merkel 2015). About the preservation of traffic data she claims that the right to unmonitored communication remained intact although internet providers are allowed

to register our data. The access to saved data in a criminal investigation would only be granted through a judge's order.

In the case of Apple's appeal to privacy protection the goal is allegedly about fighting terrorism. But what if the lines between right and wrong are not that clear another time? What if indeed the government is the one we need protection from? Like in the 1940s when the Nazis used French telephone records to track down friends of people that were already under arrest (Schneier 2001, p. 32). As I said before, we have no control whatsoever about how our information circulates in the net and a lot can change in the time span of "forever". So even if we do not live through another period of total surveillance and persecution, our children might and we do not know what we leave behind data-wise.

Kurz and Rieger also warn about the long-term risks of the online documentation of one's life. In fact, they claim that the only picture that is safe from being abused is the one that has never been taken (Kurz and Rieger 2011, p. 81). This is especially important when it comes to the trend that parents post their baby's pictures online. By that they completely deny their children's right to privacy. We should make a bigger effort to understand the intentions of major internet companies' CEOs when they talk about "giving anyone in the world the power to share anything they want with anyone" (Zuckerberg in Cocktail 2016). That is indeed to be taken literally. In the face of data collection we, the users, are nothing more (or less) than the quantifiable sum of preferences, categorized by the size of their wallet, as Kurz and Rieger put it (Kurz and Rieger 2011, p. 60).

Conclusion

With the digitization of global markets we have entered a new era of communication. We expect a lot from internet services and we take it for granted to enjoy everything for free. We should be aware that we do indeed pay for it and data is the currency. That has flicked a switch that is unlikely to reverse. We surrender our right to privacy happily and joyfully while using Facebook's online applications. We neglect our responsibility to think critically, to research and to scrutinize. We have created a world in which it is considered a comfort to be constantly confronted with ads and product placement, in which the answer that we like is preferred to the right one, in which babies are online shortly after they are born.

I am a strong believer in the internet as a common good, a free resource. Free, as in both free of charge and freely accessible. The internet offers a way to social equality like there has never been before. Unfortunately, the neoliberal world as it is produced a way to privatize that common good before it ever unfolded all its potential. But I do not want to bury that glory ideal yet. I think the internet is still one of the greatest inventions of our time. But it changes rapidly and we have to change according to it. We have to deal with legal and ethical questions along with technical ones in software engineering. And, most importantly, we have to learn to participate as active members in the online community. That does not mean playing internet trolls and post the meanest words we can think of under a random newspaper article. That means to adopt basic coding skills, to learn about data management and to read license agreements. It also means to boycott apps that become too daring. Online companies are afraid to lose customers like any other business is. Billions of internet users can provide pressure to the service providers to change their policies, if we do not permit the use of our data. Also, we can part take in think tanks and initiatives that advocate free access to the net and fight legislation that allow massive data preservation.

Even better, we can build up our own open source online communities. The open source ideal is one that has shaped the internet from the beginning. It is the one that is absolutely worth pursuing and nurturing. That way, while we want to understand the chances of big data as a resource we cannot deny the broader implications of profit-oriented corporations as the gate keepers. They are not scholarly institutes with an academic code of ethics and a dedication to research. Their interest lies only in optimizing their business model and we should not forget that when talking about online resources and the knowledge they could provide. The only way I could imagine a truly democratic and resourceful online community is an open source project were all contributors give informed consent and can participate in the research in whatever way they see fit. Through the collectivization of software development I would not be partly scared when I read about the engineering of artificial intelligence. Because the results of such research can be implemented to serve common goals. They would not only be proprietary entertainment robots which make few people insanely rich. They could be used in traffic, online, in factories, at home or wherever, to make all our lives more comfortable and enjoyable.

References

Boyd, Danah and Kate Crawford (2012). “Critical Questions for Big Data”. In: *Information, Communication & Society* 15.5, pp. 662–679.

Cocktail (2016). *Facebook 10 Year Roadmap | Mark Zuckerberg*. URL: https://www.youtube.com/watch?v=EwfNToC3LkY (visited on 09/12/2016).

Kleinz, Torsten (2012). *Wie Facebook mit Ihren Daten Geld verdient*. c't. URL: http://www.heise.de/ct/ausgabe/2012-12-Wie-Facebook-mit-Ihren-Daten-Geld-verdient-2345376.html (visited on 09/09/2016).

Kurz, Constanze and Frank Rieger (2011). *Die Datenfresser: Wie Internetfirmen und Staat sich unsere persönlichen Daten einverleiben und wie wir die Kontrolle darüber zurückerlangen*. 2nd ed. Frankfurt am Main: S. Fischer Verlag. 272 pp.

Lichtblau, Eric and Katie Benner (2016). *Apple Fights Order to Unlock San Bernardino Gunman's iPhone*. The New York Times. URL: http://www.nytimes.com/2016/02/18/technology/apple-timothy-cook-fbi-san-bernardino.html (visited on 09/12/2016).

LSE (2011). *The Filter Bubble: What The Internet Is Hiding From You [Slide-Audio]*. URL: https://www.youtube.com/watch?v=Dua_UvR5mtI (visited on 09/07/2016).

Merkel, Angela (2015). *Bundesregierung | Aktuelles | Rede von Bundeskanzlerin Merkel beim Publishers‘ Summit des Verbands Deutscher Zeitschriftenverleger (VDZ) am 2. November 2015*. URL: https://www.bundesregierung.de/Content/DE/Rede/2015/11/2015-11-03-merkel-publisher-summit.html (visited on 09/13/2016).

Pariser, Eli (2011). *The Filter Bubble: What The Internet Is Hiding From You*. Penguin UK. 246 pp.

Schneier, Bruce (2001). *"Secrets & Lies. IT-Sicherheit in einer vernetzten Welt"*. 1. Aufl. Heidelberg; Wiley: dpunkt.verlag/Wiley. 400 pp.

Sokolov, Daniel AJ (2016). *Google bittet Sie um mehr persönliche Daten*. heise online. URL: http://www.heise.de/newsticker/meldung/Google-bittet-Sie-um-mehr-persoenliche-Daten-3250642.html (visited on 09/09/2016).

Swan, Melanie (2013). “The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery”. In: *Big Data* 1.2, pp. 85–99.

Zajonc, Robert B. (1968). “Attitudinal effects of mere exposure.” In: *Journal of personality and social psychology* 9.2.