REFLECTIONS ON PRIVACY IN THE AGE OF GLOBAL ELECTRONIC DATA PROCESSING WITH A FOCUS ON DATA PROCESSING PRACTICES OF FACEBOOK

by

ANDREAS KIRCHNER[*]

In this paper I will argue that (1) individuals that are part of a community do not have full control over how information about themselves are communicated to others. Whenever we write some text messages or when we orally tell secrets to our best friend, we live with the possibility that parts and derivates of that information eventually reach unintended receivers. (2) However, there is a difference between losing control over the distribution of information and giving control to a central instance that channels communication. Facebook is a central instance that channels digital messages over a platform that is controlled, maintained and observed by Facebook. This changes the dynamics and rules of digital messaging. (3) Although Facebook presents itself as a public service to empower people and to make the world more open and connected, the service is itself closed and isolated, an experimental lab in which data flows between users are observed and categorized in order to create user profiles, summarize the profiles to target groups, and present them to advertizers and - if asked - law enforcement agencies. (4) Fortunately, social network services like Facebook are at an early stage in which users sometimes find ways to influence the evolution of the service. I summarize a recent case in which the student Max Schrems managed to put pressure on Facebook in order to change their data processing practices.

KEYWORDS
privacy, data processing, Facebook, cloud computing.

[*]   Bachelor of Arts in Philosophy, University of Vienna; Master of Science in Medical Informatics, Vienna University of Technology, Austria. E-Mail: andreas.kirchner@fragtsich.eu

1. INTRODUCTION

Facebook presents itself as a public service to empower people and to make the world more open and connected. In 2009, users protested after a change of the terms of use. The new terms made explicit that user data is stored even when the user quits Facebook. The changes were reverted and the 27-year old CEO Mark Zuckerberg says afterwards: "If [Facebook] were a country, it would be the sixth most populated country in the world. Our terms aren't just a document that protect our rights; it's the governing document for how the service is used by everyone across the world. Given its importance, we need to make sure the terms reflect the principles and values of the people using the service." (Zuckerberg 2009)

What are "the principles and values of people" that should build the basis of the terms of use? In 2010, one year later, Zuckerberg says in an interview that the social norms changed and that people are more willing to share their data publicly (Read Write Web 2010). In other words, that society gradually moves to the state of post-privacy. Post-privacy describes a society that does not attach importance to a private realm, hence it will not legally protect this realm against public use and abuse. Other CEOs of big Internet companies provide similar diagnoses. For example, Google CEO Eric Schmid when asked about Google's cooperation with governmental authorities says "If you have something you don't want anyone to know, maybe you shouldn't be doing it" (Huffington Post 2011). Already in 1999, the cofounder of Sun Microsystems said: "You have zero privacy anyway; get over it." (Turkle 2011, p. 317). If the private realm is indeed out of date, how can it be that users protest when Facebook stores and accesses data even after their user accounts are "deleted"?

In this paper I will argue that in a community an individual cannot have full control over how information about itself is communicated to others. Whenever we write some text messages, we have to live with the fact that sensitive data could reach to unintended receivers sometime. However, the loss of control over data distribution is different from an instance that is taking control over data distribution. We are not used and to a situation where data centers systematically store and restructure more data about our social and private life than we ourselves can keep in our memory.[1]

---

[1]   Although secularization made implausible an all-seeing instance like god, it seems that it never had been that much implemented, of course in form of a software that mediates and prefigures online communication. But that may be a topic on it's own.

After an initial section about Information Privacy and another about the evolution of data processing, I will argue that with the current form of electronic data processing the primary subject of privacy – the agent that "decides" which data is communicated to others – is an automatic data processing routine. Then, only secondary, we are empowered to control some of our data. The multiple data centers constitute the unconscious memory of our data and not all pop up at the user interface of Facebook. Recent findings by the austrian student Max Schrems and the report of the Irish Data Protection Commissioner exemplify the complexity of the topic and suggest that global electronic data processing is at an early stage and require watchfulness on the part of users, operators, and law institutions.

I will end by comparing the current situation with social network service with the evolving genre of tragicomedy in the 17th century and 18th century.

2. WHAT IS INFORMATION PRIVACY?

In 1890, Warren and Brandeis understood privacy as a right – the right to be let alone. He states that law must protect the "sacred precincts of private and domestic life" (Warren, Brandeis, 1890).

Approximately 120 years after Warren and Brandeis, we have to admit that our private and domestic life is not that sacred anymore and that Zuckerberg is not totally wrong with his diagnosis that the social norms have changed. On the one hand, Warren and Brandeis still capture recent fears, when we think of the wiretapping scandals in the UK. The Newspaper "News of the World" invaded privacy by hacking phones. On the other hand, it is a common and accepted practice to invade private realms of others and we know and expect that others invade ours, e.g., public blog posts, public video diaries, or reality TV.

Softening the borders between private and public life is not only negative. For example, it reduces the possibility that privacy can be used "as a shield to cover up domination, degradation and abuse of women and others" (DeCew 2008).

A more recent notion of information privacy comes from William Parent (Parent 1983). He does not see privacy as an own right but more related to a right to liberty. According to William Parent, we lose privacy when others get unauthorized access to undocumented sensitive information about us.

An undocumented information is information which is not already access-
ible in public documents.

Alan Westin gives the following definition of privacy: Privacy is an abil-
ity of a person to decide on her own when how and to what extent informa-
tion about herself is communicated to others (Westin 1967). In the follow-
ing, when I mention privacy, I mean privacy as ability.

3. THREE WAVES IN DATA PROCESSING

I consider three waves in electronic data processing (EDP). Each one af-
fected the role of EDP in the daily (working) life. My task is to make clear
the relation of social network services like Facebook with electronic data
processing. Hence, I can only roughly describe each wave using an exem-
plary device although different devices are coved by the waves and al-
though I only take up selected aspects of the wave. The aim of this section is
to better understand data processing situations and how they shaped our
habits with computers. These situations prepared the technical possibilities
and the success of social network services like Facebook.

First, the early mainframes after the Second World War usually were
isolated machines that filled a room. The task of a mainframe was pure elec-
tronic data processing to support organizational tasks to lower office costs.
It locally stored processed data, although not every result was stored inside
the machine, some were printed out immediately.

The theoretical model of the first wave of EDP is a Turing machine. Alan
Turing was inspired by people performing repetitive calculation tasks step-
by-step. He idealized the situation of a computer. By computer he meant an
individual that manipulates symbols by following rules. He formalized the
situation and showed that a fictive machine is able to perform the same
tasks.[2] I will name this kind of EDP local EDP. Local EDP is incorporated in
each of the following two waves because every electronic data procession
more or less relies on the trias input - procession – output that is formalized
in Turing machines.

In the second wave, electronic data processing and telecommunication
started to melt together. But still, at this point, local EDP was the dominant
factor. The exemplary device for the second wave is a personal computer

---

[2]   „Computing is normally done by writing certain symbols on paper. [...] The behaviour of
       the computer at any moment is determined by the symbols which he is observing, and his "
       state of mind " at that moment. [...] We may now construct a machine to do the work of this
       computer." (Turing 1936)

(PC) of the beginning 90ies of the last century. Two characteristics of those PCs are an operating system with graphical user interface and a dial-in modem to occasionally connect to the Internet. The most common metaphor to increase usability with the help of graphical user-interfaces is the desktop metaphor. In the early stages, it reminded the user of a real working place, where different objects can be placed and manipulated. This made it easier for users to interact with multiple programs. One of the programs was the browser. With the browser the user was able to send and receive data from remote servers. But before, she had to dial into the Internet via Internet Service Providers. Dialing-in was a conscious act and a relatively complex procedure. Because he shad to pay for each online minute, she had an interest in downloading necessary information to her local hard drive, go offline and utilize it later.

The introduction of telecommunication into electronic data processing initially enforced the PC as local, universal data processing unit. Browsing the Internet was only a special case of various other "offline" use cases of personal computers, e.g. games or office applications. Although remote servers processed data of the user, this happened only occasionally and within restricted timeslots. Systematic collection and representation of user data was hardly a primary goal of early-stage Internet use- or business cases.

The third wave is the most radical one, because it undermines the desktop metaphor and wrapped around the function of the local computation unit. From the point of view of the user, the working place is still a screen at home or at work. But the procession and collection of data more and more moved to remote servers. The cloud – a network of servers with central control that offers global services - gives the impression that the user is working on his private computer but in fact all data is stored and backed up in server farms of big companies. The benefit for the user is that she can access her data from an arbitrary computer, it is easy to share data with others, and she has the impression that she will never run out of disc space and performance.

In the third wave, the dominance of local data processing changed because telecommunication more and more shaped the technical circumstances as well as the user habits. Computer scientists even tried to recapture this new form of computation by customizing the theoretical model of Turing machines. They now differentiate between closed and open compu-

tation. When closed computation describes the computational power of isolated machines, open computation focuses on communication and interaction between multiple computation elements. One related model to capture open computation is the "interactive Turing machine with advice". A computer is now an agent, that interacts with other computers to produce temporary results of a global computation, i.e., "[…] a potentially endless process of interaction among components (or agents) and of components with their environment." (van Leeuwen, Wiedermann 2000, p.5). Because we focus on EDP, I name the kind of EDP that occurs in the third wave, global EDP.

The new data processing situation is best represented in a smartphone. A smartphone is a combination of a mobile computer and a mobile phone. It is supposed to help us managing our daily life and stay connected with others. Hence, we carry it with us the whole day. Two important features of a smartphone are sensors and wireless Internet: Multiple sensors are used to process local data from the environment. High-speed, wireless access to the Internet is used to access global services and information, but also to analyze local (sensor) data.

A smartphone comes to its full potential when it is used as gateway to the cloud, i.e., when the owner allows the device to delegate some of the computations to servers. One reason is that these servers have more computation power than the phone. But a more important reason is that the servers have the infrastructure to connect the owner with her friends and synchronize their data with other devices. The price of being connected to others via clouds is that the content the user shares can be used and abused by others, especially by the owners of the cloud. While in oral contexts, the audience of the speaker's proposition is limited to the people she sees around her, sharing a text message is immediately prone to abuse, because the cloud has a memory that includes the possibility that unintended receivers read the message.

4. VIRTUAL PRIVACY AND GLOBAL ELECTRONIC DATA PROCESSING

The third wave more and more replaces local with global electronic data processing. User-generated data is immediately uploaded to the cloud and associated with a user profile. The consequence is that also privacy – i.e., the ability to control how user data is communicated to others - is mediated by

the cloud. The redistribution of user data depends on the cloud, on a net-work of servers mostly owned by big companies.

Of course, it is naïve to assume that during local data processing of the second wave, the user had full control of her created content. There has been Malware, chat logs, tracing of users based on their IP address, etc. But with the shift to global EDP, it is not only the case that the user does not have full control over his data, but the primary subject of privacy changed from the user to the cloud. Hence the cloud (and the cloud operator) has more access to user-generated content than the user. Web applications like Facebook or Prezi come with default permissions on who can see your data under which circumstances. Only secondary and not necessarily[3], the cloud offers the user to restrict visibility for other users and for the user herself. This is a derived or virtual privacy that depends on the interests and the good-will of the cloud-operator - and of the limitations of the law.

The following exemplifies virtual privacy in Facebook and explicates a structural issue of global electronic data processing in Facebook and simil-ar services of the third wave. Max Schrems, an Austrian student, and his group "Europe versus Facebook" wanted to know more about the data that Facebook stores about them (Personal Data Ecosystem Consortium 2011). Schrems requested his data from Facebook and after he reminded Facebook of their obligation to provide this data, they got a CD with a 2000 pages database dump. Schrems also published step-by-step instructions on how to request personal data stored on Facebook servers.

One interesting fact that Max Schrems discovered is that Facebook does not always delete data when the user presses a "delete" button on the Face-book platform. For example, when a Facebook user deletes a message from her inbox, the message is moved to the folder "deleted messages". When the user deletes the message from the "deleted" folder, the message disap-pears from the user-interface and one might assume that the message is de-leted from the Facebook servers. The analysis of the database dump let Schrems conclude that the message is just marked as "deleted". This is a case of virtual privacy. Since not all user-related data are visible and ac-cessible to the user, she never knows which activity on Facebook (and on pages that are associated with Facebook – e.g. pages that embed the "like"-button) is stored and for how long.

---

[3]   For example, the free version of the web application Prezi only allows to create public presentation slides (Prezi 2011).

With this and other discoveries "Europe versus Facebook" submitted a document with 22 complaints to the Irish Data Protection Commissioner (IDPC) in August 2011 and an update in September 2011. In response, the IDPC did a 3-month audit of Facebook in the end of 2011. The aim of the audit was to check compliance with the EU data protection law. The results of the audit are presented in a 150-page report (Irish Data Protection Commissioner 2011).

The report of the IDPC also contains a response of Facebook on their practices in deleting messages. Since Facebook stores user data in multiple data centers it is a complex task to implement deletion mechanisms that cover all data centers and its backups. Facebook states that there might be a delay of maximal 90 days until a post is fully deleted from all Facebook severs. Moreover, Facebook can only fully delete the message, when the sender and the receiver have deleted the message. Facebook "states that its policy and practice is to delete a message after the last person user deletes the message". (Irish Data Protection Commissioner 2011, p.73) Note, that the last statement was not verified during the audit. The Commissioner will check the statement during the review.

In summary, taking a closer look into the way how Facebook processes data helps to better understand the situation. On the one hand, the actions of Max Schrems focused the excitement by providing concrete complaints. Moreover, the report of the IDPC calmed down the excitement a bit. Facebook does not appear any more as the unstoppable data-collecting beast. It has to comply with European law and committed to immediately start action to do so. Sober observations and efforts of a small group led to an intensive company audit and helped to get more control and knowledge of the stored data.

On the other hand, this example confirms the shift from local, uni-centric data processing of the first and second wave of EDP to global, multi-centric data processing that makes it difficult to control data movements, even for operators. Several data centers located in different countries with different data protection laws and different backup strategies pose new challenges for lawyers, computer scientists and data protectionists.

5. INCENTIVES TO COLLECT AND PROVIDE USER DATA

Collecting and storing so much data is a significant technical and financial expenditure. So why does Facebook offer their social network service for

free? A simple answer of Constanze Kurz – a spokeswomen of the German Chaos Computer Club – is, that users are not the primary customers. They – and especially their activities inside the platform – are the product. The real customers are advertisement companies. (Deutschlandradio Kultur 2011).

From the economic perspective, the incentive of Social Network Services to store so much user data can be better understood with the autonomous business model. The autonomous business model betrays the user as a prosumer. The user produces and consumes data via an online platform. The data is then stored in the cloud. Facebook analyzes this data with data mining methods and creates user profiles to categorize users based on their interests and activities. Facebook then offers advertisers to target different interest groups. As a side effect, it also grants (and is legally obliged to grant) access to law enforcement agencies like the FBI if they request data.

A tutorial that describes the autonomous business model notes that companies that use that business model basically gain money by doing nothing – at least they do not provide content/goods. They do not provide any content except an initial content as a showcase (Flor 2001). This would not be possible in traditional or "offline business".

What are the user motivations to generate content inside Social network services that can be sold to advertisers? To answer again with Constanze Kurz: Facebook is designed to have a high "stickiness factor" (Deutschlandradio Kultur 2011). Those new forms of synchronous and asynchronous communication are interesting, fascinating and also useful for people.

Social network services sometimes even overbid other excitements like watching TV, because you can always say that you are investing time in your social network. "It is not just fiction, it is relevant to your social life", can be a direct excuse for staring on the screen. In addition to that, Sherry Turkle says, that this kind of interaction promises safety and acceptance. Whatever you do, you are in distance to the people you feel close to (Turkle 2011, p.321).

One common political reason to use Facebook is described by Waddick Doyle and Matthew Fraser: "Since Web 2.0 networks diffuse power away from institutions and towards people, social networking sites are lauded as effective platforms for promoting a genuinely bottom-up expression of citizen sovereignty." (Doyle, Fraser 2010, p. 226)

In fact, Web 2.0 in general does not diffuse power away from institutions and towards people. Social network sites only rearrange the power and give

users new opportunities to pose and expose. The formulation "If Facebook were a country" does not only relate the number of Facebook members to the number of citizens in a country. It also suggests a similarity between a social network service (with "governing" terms of use) and a country that controls its citizens based on a common legislation that people accepted but sometimes forget or ignore.

For the first time in history we have an efficient communication to globally organize ourselves to fight against established regimes. But when we take into account the technical and organizational aspects of the IT company Facebook, the list of complaints from Max Schrems and the report of the IDPC, then it is inadequate to say that Facebook only diffuses power away from institutions. In contrast to other forms of Internet communication like blogs, Facebook is a centralized architecture, like Twitter and Google+. I cannot install Facebook on my personal computer or on my own web server, I need to use my computer as a gateway to Facebook. The advantage of a cloud services is that we do not need to care about most of the technical details (like security, long-term backup, uptime, spam, etc.). On the other hand, one effect is that millions of people do not only lose their ability to control their data, they give it to a significant part in the hands of a private institution. Moreover, law enforcement agencies tap the wealth of categorized user behavior. The virtual privacy settings of Facebook are ineffective to restrict the influence of such institutions. In contrast to a house search, the user might not even know that law enforcement agencies have gained access to personal data, because the cloud operators decide whether to provide them access or demand a proper search warrant.

6. TRAGICOMEDY

The last section aims to get distance to exciting developments about electronic data processing by looking at a different topic in history. Tragicomedy is a genre in fictional work that rearranges and interrelates comedy and tragedy to a single unit which makes it difficult to separate them - although you can analyze both aspects separately. This seems paradox at the first glance. Karl Guthke tried to capture this genre theoretically (Guthke 1966). He starts with the observation that Tragicomedy is a modern phenomenon – a phenomenon of the 19th and 20th century. But before it established as a modern genre, the mixture of tragic and comic aspects was a main controversial debate about "good taste" in neo-classicismic theory of the 17th and

18th century. At this time, the traditional rules of established authorities in literary theory became outdated. The traditional rules forbade any mixture of comedy and tragedy. Some neo-classicists discovered that the "aesthetic sense" of the reader enjoys this mixture. Others – for example those who supported the rule of imitation of nature, rejected any mixture of opposites because nature was considered as harmonic and ordered. So, the precursors of modern tragicomedy were called "barbaric", "destructive", etc. Others like the english poet John Dryden judged more carefully and did not condemn the thing, but their manner of doing it. Then, Dr. Samuel Johnson (the most cited person in the 18th century) related tragicomedy with personal experience. He stated that the audience already knows the mixture of tragic and comic affections from daily life. Based on this discussion the modern tragicomedy and systematic theories about this genre arose in the 19th and 20th century.

How does this controversy relate to privacy? In my opinion, Internet communication (or in general: Telecommunication) is a new genre in inter-personal communication. It has the touch of a private conversation that we know from oral communication. At the same time, it can be easily replicated and redistributed like written communication. This confuses our well-established domains of private and public realms. Although we can analyze both aspects separately, Telecommunication itself is private and public at the same time. It will take a while for us to find stable ways of coping with this new genre.

Facebook is one variant of this genre. Zuckerberg says, the mission of Facebook is to make the world more open and connected. I think that this is an adequate description, when we add the following: The system that actually makes the world more open and connected is itself closed and isolated, a walled garden. The Facebook machinery is to a high degree private. We should not forget, what the Irish Data Protection Commissioner only mentioned in passing: "FB-I [Facebook Ireland, A.K.] is the "data controller" in respect of the personal data of these users" (Irish Data Protection Commissioner 2011, p.3). This accumulation of power can be abused. "But in fact we are in early days. There is time to make the corrections" (Turkle 2011, p. 364). When we reconsider how the group "Europe versus Facebook" consisting of only a few students figured out to cause a company audit, we already have an example of how to correct developments.

REFERENCES

[1] Zuckerberg, M. 2009, Update on Terms
https://www.facebook.com/blog.php?post=54746167130
[Accessed Jan 03 2012]

[2] Huffington Post 2011, Google CEO on Privacy
http://www.huffingtonpost.com/2009/12/07/google-ceo-on-privacy-
if_n_383105.html [Accessed Jan 03 2012]

[3] ReadWriteWeb 2010, Facebooks Zuckerberg Says The Age of Privacy Is
Over https://www.readwriteweb.com/archives/facebooks_zuckerberg_says
_the_age_of_privacy_is_ov.php [Accessed Jan 03 2012]

[4] DeCew, J. 2008, Privacy, fall 2008 edition, The Stanford Encyclopedia of
Philosophy http://plato.stanford.edu/entries/privacy [Accessed Jan 04 2012]

[5] Warren, Brandeis, The Right to Privacy, Harvard Law Review 193,
http://groups.csail.mit.edu/mac/classes/6.805/articles/privacy/Privacy_brand
_warr2.html [Accessed Jan 15 2012]

[6] Parent, W. 1983, Privacy, Morality and the Law, Philosophy and Public
Affairs 12: 269-88

[7] Westin, A. 1967, Privacy and Freedom, Atheneum, New York

[8] Turing, A. 1936, On Computable Numbers, with an Application to the
Entscheidungsproblem, Proceedings of the London Mathematical Society
42:230-265

[9] Van Leeuwen, J., Wiedermann, J., 2000, The Turing Machine Paradigm
in Contemporary Computing, Mathematics Unlimited, Springer, 1139-1155

[10] Prezi, 2011, Help for newbie!, http://community.prezi.com/prezi/to-
pics/help_for_newbie#reply_6149535 [Accessed Jan 14 2012]

[11] Personal Data Ecosystem Consortium, 2011, Austrian students report
Facebook to Ireland's Data Protection Commissioner, http://personaldatae-
cosystem.org/2011/09/austrian-students-report-facebook-to-irelands-data-
protection-commissioner/ [Accessed Jan 15 2012]

[12] Irish Data Protection Commissioner, 2011, Report of Data Protection
Audit of Facebook Ireland Published, http://dataprotection.ie/viewdoc.asp?
DocID=1182&m=f [Accessed Jan 15 2012]

[13] Deutschlandradio Kultur, 2011, Benutzer werden eigentlich kategoris-
iert, http://www.dradio.de/dkultur/sendungen/tacheles/1596320/ [Accessed
Jan 15 2012]

[14] Flor, Nick. V, 2001, The Business Model Approach to Web Site Design,
http://www.informit.com/articles/article.aspx?p=20882
[Accessed Jan 15, 2012]

[15] Turkle, Sherry, 2011, Alone Together: Why We Expect More From Technology And Less From Each Other, Basic Books, New York

[16] Doyle, Waddick, Fraser, Matthew, 2010, Surveillance and Power, Facebook and Philosophy: What's on Your Mind? (Popular Culture and Philosophy), ed. D.E. Wittkower, Open Court

[17] Guthke, Karl, 1966, Modern Tragicomedy: An Investigation Into The Nature of The Genre, Random House, New York