THE POWER OVER PRIVATE INFORMATION IN BIG DATA-SOCIETY

Power Structures of User-generated Data Manifested by Privacy and Data Policies

Charlotte Högberg

Master's thesis (30 ECTS) in Library and Information Science for Master's degree in Master of Arts in Archival studies, Library & Information studies and Museum studies (ALM), Lund University
Supervisor: Sara Kjellberg
Year: 2015

Title

THE POWER OVER PRIVATE INFORMATION IN BIG DATA-SOCIETY: Power Structures of User-generated Data Manifested by Privacy and Data Policies

Abstract

The starting point of this thesis is the managing of user-generated data in the online ecosystem and expanding development of big data. Many are worried that companies and authorities are invading their online privacy, and the lack of control by the provider of data, the citizens, can be considered one of our time's most pressing civil rights issues. At the same time, media and information literacy become more and more important for the ability to actively be part of society. Libraries have an educational role to gain awareness of information issues, which includes privacy issues. The aim of this study is to *investigate the power structures of privacy, ownership, gathering, store and use, of user-generated data, through the discourses manifested by privacy and data policies of social media services.* This is done by deploying a theoretical framework of power and language with *critical discourse analysis*, CDA, and of mechanisms of privacy with *communication privacy management,* CPM, theory, complemented by a discursive understanding of power and normative manifestation in online interfaces. Methodologically the study is conducted by a critical discourse analysis of the privacy and data policies of Facebook, Twitter, Instagram, Google, Youtube, Tumblr, Pinterest, Snapchat, Reddit, Linkedin and Ello. An interface analysis is also conducted on the same social media services' mobile phone applications and websites, pre and post login. By this, different discourses are identified. The companies claim that the users' privacy is something valuable and important but this is not mirrored by the interfaces, where links to privacy policies mainly are placed in the bottom of pages and menus. In the policies privacy is constructed as possession, claiming to belong to, and be controlled by, the user. However, later statements contest this by manifesting great restrictions on both ownership and control. At the same time, the language of the policies is used to portray the user as responsible for all of the services' practices. The policies of Reddit and Ello constitute exceptions in some respects and also express discursive struggle. In conclusion, this study shows that power in the policies is manifested by uncertainties, the users' lack of control and influence and the social media companies' lack of transparency.

Keywords

Library and Information Studies, Library and Information Science, Privacy, User-generated Data, Power, Social Media, Critical Discourse Analysis, Communication Privacy Management Theory, Interface Analysis, LIS, ALM.

Biblioteks- och informationsvetenskap, Integritet, Personlig integritet, Användargenererad data, Makt, Sociala medier, Kritisk diskursanalys, Gränssnittsanalys.

CONTENTS

1. Introduction

During an information management conference in 2013, the media professor Siva Vaidhyanathan got the question whether there was something one could do to prevent having one's data gathered online. Vaidhyanathan's answer illustrates that this is not an individual issue; it is a social civil rights issue:

> Sure, there are simple things that you could do, that I could do, that a couple hundred people in this room could do. But that does not help my mom, and that does not help all the folks we work for and work with. When we are concerned about privacy, it cannot just be about our privacy; it has to be about everyone's privacy. Civil liberties are not something that we have the ability to trade away willingly, because they do not belong to us; they belong to all of us.
>
> *Vaidhyanathan & Bulock, 2014, p. 62.*

Ownership and control in the social media era of today do not only pose dilemmas regarding the copyright to the images that users publish on Instagram; the more critical dilemma is the control over information about the users, such as their name, birthdate, behavior, relations, contacts, devices, interests, search history, clicks, phone records and the content of their emails. According to a recent survey, *Svenskarna och internet 2014* by Findahl (2014), one in four Swedish citizens are worried that large companies, such as Google and Facebook, are invading their online privacy and one in five are worried that authorities do it as well. The same survey shows that 16 percent of the Swedish Internet users consider, and accept, that privacy is something that no longer exists. This is intertwined with the development of big data - a development where information about individuals can be used without them being aware of, or in control of, the gathering and use of their data.

Big data-analyses open up a whole range of new possibilities for research and development (Lane, Stodden, Bender, & Nissenbaum, 2014). The analyses can consist of information, gathered through Facebook, Twitter, Instagram, YouTube, health records or official documents. Analyses of big data can be used to attain social benefits: to prevent illnesses, handle accidents and catastrophes or to find out where resources are most needed in a city. But this comes with a price - invasion of privacy, and could entail surveillance of individuals and registration of individuals' views and political statements (Lane et al., 2014). Big data is mostly discussed as a phenomenon for business intelligence, personalized advertising or as of methodological interest for research. It is the "talk of the future", by some considered hype, but large investments are made. The European Union is for example investing billions of euros in development and research on big data the years to come (Wallström, 2014).

In terms of the larger picture the use of different sorts of user-generated data in society and the lack of control by the provider of data, the citizens, can be considered

one of our time's most pressing civil rights issues (Croll, 2012). While writing this, Facebook could be facing a class action lawsuit filed by 25,000 Austrian Facebook users that claim that the company has been invading their privacy, illegally tracking their data and cooperated with US National Security Agency on surveillance (Gibbs, 2015). The law student that initiated the lawsuit, Max Schrems, says:

> Basically we are asking Facebook to stop mass surveillance, to (have) a proper privacy policy that people can understand, but also to stop collecting data of people that are not even Facebook users.
>
> *Gibbs, 2015.*

Sundin and Rivano Eckerdal write that we as citizens more and more frequently have to independently search and value information. Nowadays this is an important part of people's ability to actively be part of society and while the Internet is global, the access to it and the users' competence is not equally distributed (Sundin & Rivano Eckerdal, 2014, pp. 9-11). When discussing the concept of information literacy, Limberg, Sundin and Talja (2012) state that literacy do not only encompasses the ability to read and write, but also the ability to understand and interpret texts and statements, especially when faced with contradictory messages. They also stress the empowering nature of literacy:

> literacy does not only transform individuals but is also the condition for individuals' power to transform society. Literacy therefore extends from a mechanical skill to the ability to think critically and challenge dominant ideologies.
>
> *Limberg et al., 2012, p. 98.*

It is vital to discuss privacy issues in the library sphere, since libraries have an educational role of media and information literacy as well as serve to protect the ability to freely search for information without surveillance. Libraries hold the integrity of the users high, and as said by Siva Vaidhyanathan:

> It's extremely important that the library community […] become[s] the most vocal citizens on these matters. Librarians understand that a key element of intellectual freedom is the ability not to be interrogated about your curiosity.
>
> *Vaidhyanathan & Bulock, 2014, p. 61.*

Similar conclusions were made by Barbara Jones in her talk on patron's privacy at the Swedish library conference *Biblioteksdagarna* 2014 (Svensk biblioteksförening, 2014). Social media as part of digital literacy that patrons could be helped to engage in, and the online management of personal information and user-generated data, becomes an issue for everyone working with information. For those set out to protect privacy it is important to raise awareness of discourses concerning it.

One concern is the ownership of the users' personal information. For example, due to signed user conditions, Facebook can use data about individuals as commercial goods to make money by "sharing it" with advertising agencies (Buchanan 2011). Similarly, Google uses personal data to shape their advertisement and services, in order to make profit. Social media companies also release personal information upon governmental requests. During 2013 Facebook received requests to reveal information about over 170 user profiles in Sweden (Facebook, 2015d). During the same period information

about 90 users/accounts were requested from Google by the Swedish government (Google, 2015f). Clearly there are asymmetrical power relations between the creators of data and those in charge of the services it is collected through, by whom the personal information is passed on to third parties. boyd and Crawford (2012) writes that these inequalities of power are written into the system and therefore producing class-based structures. Manovich (2011) have formulated that the big data society divide people and organizations in mainly three categories:

> Those who create data (both consciously and by leaving digital footprints), those who have the means to collect it, and those who have expertise to analyze it. The first group includes pretty much everybody in the world who is using the web and/or mobile phones; the second group is smaller; and the third group is much smaller still. We can refer to these three groups as new "data-classes" of our "big data society".
>
> *Manovich, 2011, pp. 10-11.*

I want to further investigate how these power relations are manifested, which includes several aspects. The most apparent is the control and ownership of the user-generated data, but also the notion of privacy, concerning people's personal information and user-generated data. This means what privacy is considered to consist of, what it is needed for and who has right to it. These questions can be approached by analyzing the discourses of policies regarding user-generated data. Since social media are mainly consisting of user-generated data and are widely used, the language of privacy and data policies of social media constitutes the empirical foundation of this study. To fully grasp the power structures that shape the privacy policies I will also need to consider the context that they are part of. In the most direct way; the online interfaces in which the policies can be reached by, and correspond with, their intended audience; the users. Conclusively, the policies and their online contexts need to be discussed in the realm of big data.

1.1 Research Aim

The general purpose of this thesis is to contribute to the understanding of how power structures can affect information practices. There is a call for more knowledge concerning how power relations can influence information behavior and practice within the field of library and information studies (Olsson, 2010, Haider, 2008, Olson, 2002). This also concerns information management, and it is especially important with the expanding use of digital tools like social media and online profiles, as well as the development of big data and its possible impact on society. Besides academic and professional relevance, this study is intended to contribute to increased understanding of power structures influence on online privacy, personal information and big data, which today is relevant for society as a whole.

Libraries have an educational role regarding information issues, which privacy issues are part of, as well as a code of ethics that stresses the protection of citizens' privacy. It is relevant to relate the notion of privacy and power over user-generated data to the library sphere, and as situated in library and information science, since libraries has a role to support media and information literacy. As previously mentioned, many Swedes are concerned about their privacy online. However, 52 percent believe that they can control their own privacy online, according to *Svenskarna och internet 2014*,

but when asked what measures they take to control it, not many actions are mentioned (Findahl, 2014). The most common is to delete cookies from the web browser from time to time, something that mainly aims to speed up the computer. The large majority, 81 percent, does not take any measures. A third of the women claim that they cannot answer the questions since they are not sure what the questions on technical measures refer to (Findahl, 2014). Some researchers (e.g. Debatin 2011) stress the need for public education on *privacy literacy*, where libraries have a role to play. At the same time Zimmer (2013) also emphasizes the need for discussion and policies regarding privacy, and education for information professionals, within the LIS-field. Altogether, this calls for a greater focus on privacy issues within media and information literacy education, and hence on privacy issues within library and information science.

Language affects how we perceive reality and has real effects; language is also affected by non-linguistic factors of reality. By using discourse analysis to study how privacy and data policies of social media services are expressed, and incorporate it with an analysis of how the policies can be found in their online interfaces, this study can contribute to the understanding of the power mechanisms, manifestations of power and ideological workings of discourses of privacy and user-generated data. By using theoretical perspectives on privacy this study contributes to the understanding of mechanisms of privacy, the notion of the concept and the need for personal information to remain private, as well as what constitutes violations against privacy and how mechanisms to exercise power work within the concept.

In conclusion, the research aim of this thesis is to: *Investigate the power structures of privacy, ownership, gathering, store and use, of user-generated data, through the discourses manifested by privacy and data policies of social media services.*

1.2 Research Questions

The research questions I will answer in this thesis are the following:

1. How is the concept of privacy constructed by the policies of social media services?
2. How are the owner conditions, in terms of collecting, storing and usage, of the user-generated data mediated?
3. How is the user's room for action depicted?
4. Which power relations can be detected in the policies and how they can be found in online interfaces, and how can these power relations be understood in relation to the socio-political context of big data?

1.3 Background

In this chapter I work out the landscape that this study arises from. This includes descriptions of the concepts discourse, power and privacy as well as of the phenomena social media, information, user-generated data and policies. Lastly, I discuss how the right to online privacy relates to the library sphere.

1.3.1 Discourse and Power

By using the concept of discourse, one has to acknowledge and accept a set of philosophical stipulations about societies, the world and reality in general: that perceptions are changeable and vary between societies and time periods and that the way we, by the use of language, classify and interpret the world has real consequences (Boréus, 2013b, pp. 150-151). The person most associated with the concept of discourse is Foucault. From a foucauldian perspective discourse can be defined as changeable set of rules which legitimize certain knowledge and not other, and determine whom is given authority of expression (Bergström & Boréus, 2008, p. 309). As mentioned by Boréus (2013b, p.151), later modifications to discourse analysis emphasize that the use of language and social practices affect how we interpret the world, but that the world itself also affects how we act and use language. This is the viewpoint that will be the base for this study.

Discourse analysis highlights the societal context that texts are a part of, and this contextualization is also an important factor for the analysis. Fairclough (2001, pp.16-19) describes discourse as language in the form of social practice. He argues that language is part of society and is a social process, conditioned by other non-linguistic parts of society: "Language is a part of society; linguistic phenomena are social phenomena of a special sort, and social phenomena are (in part) linguistic phenomena" (Fairclough, 2001, p. 19). The discursive notion of power is a broad definition of power as in Foucault's statement that power exists in all social relations. The core concerns are what coercive norms are created by discourse, rather than individual actors and their underlying motifs (Bergström & Boréus, 2008, p. 328). Power is exercised in different forms, and Fairclough (2001, pp. 3, 27-28) makes an important distinction between the two broadly defined ways in which power can be exercised and kept, through *coercion* and through *consent*. Power through coercion is exercised by such as physical violence, while power through consent is exercised by winning others' acceptance for one to possess and exercise power. Fairclough states that the main mechanism behind the method of ruling by consent is ideology, which plays an increasingly important part: "the exercise of power, in modern society, is increasingly achieved through ideology, and more particularly through the ideological workings of language" (Fairclough 2001, p. 2). Discourse and power are important concepts that will be further explored in the theoretical framework of this thesis.

1.3.2 Privacy

According to the Oxford English Dictionary privacy is:

> a state in which one is not observed or disturbed by other people: *she returned to the privacy of her own home*. the state of being free from public attention: *a law to restrict newspapers' freedom to invade people's privacy*.
>
> *Oxford dictionary of English, 2010.*

What the concept of privacy consists of has been explained in many different ways. Even among scholars interested in privacy matters, a definition of the concept has not been consolidated (Margulis, 2011; Vasalou, Joinson, & Houghton, 2014). Margulis (2011, pp. 14-16) calls privacy an elusive, elastic and psychological concept that includes a variety of philosophical, legal, behavioral and everyday definitions. He argues that there are disagreements regarding what to include in the concept and if the concept only concerns behaviors that are considered morally neutral or respectable, which is the most common perspective, or if it also can include illegitimate activities. According to Margulis (2011, pp. 14-16) there is no agreement on within which philosophical frame privacy should be defined. Vasalou, Joinson and Houghton (2014, p.2) argue that this lack of uniformity can be problematic when it comes to developing technology or policies that take privacy risks into concern. By a vast number of surveys they have analyzed what features people include in the concept of privacy and the results supported existing theoretical privacy definitions. Frequent features mentioned in the surveys were *secrets, being alone/without company, personal space* and *right to entitlement* but also behavior regulation components as *having choice, protecting personal information, having control over one's information* (Vasalou, Joinson & Houghton, 2014, pp. 9-11). This thesis will not be limited to any single one of these fixed definitions of the concept of privacy, since the aim is to investigate discursive power structures of privacy and investigate how the concept of privacy is constructed by the policies.

Privacy is situated as collective information practice, as well as discursive practice, by Dourish and Anderson (2006). With information practice they allude to how we share, withhold and manage information, how we interpret such acts and how we deploy them in social interaction (Dourish & Anderson, 2006, p. 335). To approach privacy as a discursive phenomenon entails analyzing how the notion of privacy is "used to categorize activities, events, and settings, separating acceptable (secure) actions from unacceptable (insecure) ones" (ibid, pp. 328-329).

1.3.3 Social Media

> Social media is predominantly a corporate-state-power phenomenon, a force field in itself, in which powerful corporate and state interests are present and meet, as evidenced by the existence of a surveillance-industrial complex (PRISM) that controls social media communication and is constituted by a collaboration of social media and Internet companies, secret services and private security companies.
>
> *Trottier & Fuchs, 2015, p. 34.*

64 percent of the Swedish citizens use social media (Findahl 2014). Social media is a broad concept emanating from the question of what it means to be social. Trottier and Fuchs (2015) argue that three social information processes can be identified to clarify the notion of social media: cognition, communication and cooperation. They refer to forms that mostly support cognition, such as newspapers' websites, as social media type 1, forms that mostly support communication, such as e-mail, as social media type 2 and forms that mostly support community building and collaborative work as social media type 3 (ibid, pp. 4-5). The importance of social media type 3 has increased, due to the rise of social networking sites (Facebook), wikis (Wikipedia)

and microblogs (Twitter), argue Trottier and Fuchs (2015, pp. 5-7). However, they describe Facebook as a type of social media that at the same time supports cognition and communication/networking, as well as cooperation: "therefore a lot of personal and social data about users is generated" (ibid, p. 7). Furthermore, Trottier and Fuchs (2015, p. 34) argue that social media are "spaces of complex manifestations of power, counter-power and power contradictions."

1.3.4 Information and User-generated Data

The concept of information is ambiguous and ascribed a variety of meanings, as discussed by Buckland (1991, pp. 3-4). He identifies three principal uses of the word information in an attempt to identify how the term is used. These are:

- Information-as-process: information as the act of informing and communicating knowledge
- Information-as-knowledge: information as the knowledge communicated in the process
- Information-as-thing: information as objects such as data and documents, objects that are informative

Information-as-knowledge is intangible; it cannot be touched or directly measured. As well as beliefs and opinions, knowledge is personal, subjective and conceptual. To communicate these phenomena "they have to be expressed, described, or represented in some physical way, as a mark, signal, text or communication" (Buckland, 1991, p. 4). This makes something that is information-as-knowledge into something that is information-as-thing, writes Buckland.

The certain type of information that this thesis will discuss is user-generated data. User-generated data is a versatile concept consisting not only of the content intentionally created and uploaded by a user, such as pictures on Instagram or status updates on Facebook. It also consists of personal information knowingly provided by the user, such as e-mail address and phone number, as well as personal information unintentionally provided, such as relations and user-activity gathered through the service. These can be activities performed on the service, such as clicks and searches, but also activates performed outside of the service. As Andrejevic (2015) states regarding Facebook:

> Facebook's entire business model […] is built around the information provided by users: not just the number of times they click on a "like" button or the network of "friends" they link to, but the minute details about which websites they visit, what they buy, what type of information they read, how often, when, and where, and the growing array of detailed information about behavior, preferences, activities, and so on that the platform is able to capture.
>
> *Andrejevic, 2015, p. 8.*

1.3.5 Policies

There are different types of texts, which can be divided into different genres. Texts are produced and consumed in differing contexts and there are various conventions regarding how texts should be formed, read and used (Boréus, 2013a, pp. 132-133). Texts that are produced in a certain genre have a certain function and a special form, such as language and structure. Boréus (2013a, pp. 132-133) acknowledges the

importance of awareness regarding which genre texts are a part of and what it entails, when analyzing texts.

The texts I choose to analyze to approach the question of power structures regarding user-generated data are the policies that deal with how this kind of data are managed by the social media companies. A policy can be described as a set of rules or guidelines to follow. Companies and others managing data are required to have these kinds of policies since "informed consent" by the user are a legal requirement by the EU data protection directive to be allowed to collect data (e.g. Bechmann 2014). The same is stated by the U.S. Fair Information Practice Principles, which demands that data subjects should be given notice on how and by whom information is collected, used and shared as well as if the information collection is voluntary or required (e.g. Barocas & Nissenbaum, 2014, pp. 56-58). Due to state law in California (where many of these businesses resides) every website or online service that collects personal information about individuals are required to post its privacy policy on its website. This policy should meet certain requirements listed in The California Online Privacy Protection Act of 2003 (amended in 2013), such as:

- Identify which categories of personally identifiable information that are collected and the categories of third parties it is shared with
- If there is a process for users to review and request changes to the personal information a description of that process should be provided
- Describe how changes of the privacy policy are notified
- Disclose how the operator respond to Web browser's "do not track" signals and other mechanism which provide alternations to the collection of personally identifiable information (State of California, 2003).

1.3.6 Libraries and the Right to Online Privacy

> No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor to attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks.
>
> *United Nations, 1948.*

The right to privacy is a basic human right, according to article 12 in the United Nations Universal Declaration of Human rights (1948). The American Library Association, ALA, has initiated the annual "Choose Privacy Week", which aims to initiate a conversation about privacy rights in a digital age. The goal of the campaign is to give libraries tools to educate and engage users and by that give citizens resources to "think critically and make more informed choices about their privacy." (American Library Association, 2015a). On the question why libraries should engage in this, ALA answers:

> Because the freedom to read and receive ideas anonymously is at the heart of individual liberty in a democracy. Librarians defend that freedom every day. Libraries are information hubs for their communities. They are also natural centers for learning and talking about information issues… including privacy.
>
> *American Library Association, 2015b.*

This highlights privacy issues as part of libraries' support for media and information literacy. Sundin and Rivano Eckerdal (2014, p. 15) emphasize the increasing educational role of libraries, for people to turn to for help on how to gain knowledge and competence to navigate today's informational landscape. They argue that media and information literacy relates to the democratic function of libraries, as libraries constitute a supposition for the democratic society by offering free access to information and knowledge. As more of the media landscape moves from print to digital the librarians educational role becomes more important, since access is not just a technical issue but also dependent on recourses and abilities. The ability to understand how information is produced online and the ecosystem of the web as well as to understand and make use of new digital media and be able to publish content in participatory media, is vital aspects of media and information literacy, writes Sundin and Rivano Eckerdal (2014, pp. 16-21).

When The International Federation of Library Associations and Institutions, IFLA, defines trends that will shape future of the information environment, one of them is: "The boundaries of privacy and data protection will be redefined" (IFLA, 2013). The report predicts that as data sets held by government and companies grow larger and methods of monitoring and filtering data becomes easier and cheaper, the profiling of individuals will become more advanced. This can result in: "serious consequences for individual privacy and trust in the online world" (IFLA, 2013, p. 12). Libraries have been considered "safe places" regarding privacy, but the report questions whether this will be possible to sustain. Libraries' role to protect digital privacy is emphasized by Gressel (2014, pp. 138-139), who questions whether libraries are doing enough. As libraries try to stay relevant by keeping up with technology they often open up to privacy threats; "libraries should be reevaluating their privacy policies and figuring out the best ways to maintaining privacy in the modern world" (Gressel, 2014, pp. 139-140). He emphasizes that privacy is a responsibility for libraries also regarding services outside of the library scope (such as Facebook), which can be acknowledge by raising awareness of privacy among patrons and by protesting breaches of citizen privacy (ibid, pp. 139-141). Views regarding libraries and their responsibilities can vary between countries with differing laws and regulations, though this thesis discuss libraries on a general level.

1.4 Disposition

Chapter one has been an introduction to, and formed the background of, this thesis. In chapter two I describe previous research, which is followed by the theoretical framework presented in chapter three. Chapter four consists of method and methodological aspects of this study. The results are presented and analyzed in chapter five, called Description. Chapter six, Interpretation, and chapter seven, Explanation, includes further analysis and discussion in relation to previous research and my theoretical framework, as well as conclusions and suggestions for further research.

2. Previous Research

In this chapter I examine previous research that is relevant for this study and that can contribute to my analysis. First research on power and discourse within library and information studies, henceforth LIS, is elaborated. This is followed by research on privacy within LIS. To cover a wider scope and reach a more comprehensive outlook, this chapter also includes research on power, privacy and policies within other fields of research.

2.1 Power and Discourse in LIS

Power has been investigated within LIS in terms of knowledge organization. Since power and language is a central theme in my thesis it is fruitful to incorporate previous understandings of how language is used to express and practice power. Olson (2002) reveals power structures in library classification systems by analyzing how they assign names and labels upon subjects and thereby shape perceptions and society. She states: ”I choose the word ”naming” because it connotes the power of controlling subject representation and, therefore, access” (Olson, 2002, p. 4). The action of bestowing a name is creating an identity and this, Olson writes, is a way of structuring reality by imposing the namer’s viewpoint of the world. She further argues that: “Naming information […] is not simply representation of information, but is also the construction of that information” (Olson, 2002, p. 6). By examining foundational texts of library cataloguing, current standards and their canonical applications she finds universal vocabularies with rigid ideas of structure and hierarchal classifications, which results in marginalization and exclusion. Subject representations are enforced by discourses of authority that uphold conformity and Olson finds that “Other” women are consistently marginalized and excluded through the naming. Language use as function of power, labeling, and constructions of concepts is also objects for this study.

Discourse within LIS has been investigated by Jutta Haider, in *Open Access and Closed Discourses: Constructing Open Access as a “Development” Issue* (2008). By examining policies, among others sources, she analyzes open access in terms of discourses discussing the “developed world” and the “developing world”. She finds that the notion of the “developing world” is constructed in line with mainstream development thought, within two discourse systems: one of science in which “Western” science is viewed as universal, global knowledge, and one of development as modernization. While the debate promotes and articulates openness it does so within closed discourses. Open access is constructed as a problem of development as well as its solution, and as a measure of progress in the bigger idea of development as an evolutionary progress. Haider finds that the discourses propose binary positions

with the world divided in a more developed and a less developed world, where the knowledge of the first is more valued. The discourses also reveal technological determinism and views upon information as something that goes from sender to receiver, by using words as "bridges" and "flows".

Haider (2008, pp. 29-30) argues that discourse analysis, and its understanding of power, is becoming more established within the field of LIS. She writes that discourse analysis mainly has been deployed in terms of general theoretical consideration, outlining possible approaches for discourse analysis in the field of LIS, to analyze interviews or, as in my case, to analyze *naturally* occurring documents and discourses, that are not the product of an interview but occur prior to the analysis.

Like discussed in the Introduction, media and information literacy has become vital to be part of the contemporary society (Sundin & Rivano Eckerdal 2014). How information literacy relates to discourse analysis has been reviewed by Limberg, Sundin and Talja (2012). They state that within librarianship, information literacy appears as an object of teaching and literacy as an outcome of learning. Further, they define information literacy as the ability to search, select and critically evaluate information, but they also acknowledge information literacy as ways of learning and as skills to manage and handle information (ibid, pp. 94-96). Limberg et al (2012, p. 98) states that the development of information and communication technologies and new media calls for a wide definition of literacy, often referred to as digital literacy and media literacy. Limberg et al argue that discourse analysis is important for mainly two reasons:

> First, when different conceptions of a specific issue are brought into view, the most self-evident and powerful viewpoints are destabilized. They tend to lose some of their credibility and status as objective truths; we come to realize that more than one truth exists. Second, analyzing variability in ways of conceptualizing the nature of a specific issue or phenomenon (such as literacy or information) opens up new viewpoints and promotes novel understandings concerning the topic at hand.
>
> *Limberg et al., 2012, p. 114.*

To apply discourse analytical perspectives on information literacy, they argue, is to investigate the meaning that is ascribed to information competence and practices, and the discourses of information literacy. However, they also state that other discourse analytic studies within LIS and other fields are relevant for information literacy research "in that they discuss conceptions of the nature of information, information needs, and information and communication technologies" (Limberg et al., 2012, p. 110). Limberg et al finds that most discourse analytic studies within LIS concerns how texts describes roles and competences of information users, where users are portrayed as in need and as the "information poor". Other studies shows different disciplinary discourses regarding critical thinking, often meaning reliability and credibility of information sources, by such as the authority and status of the creator, which represents a positivist philosophical orientation. In opposition to this an approach is posed, which instead focuses on language use and how texts are crafted to achieve effects and justify positions, as well as how texts contribute to (re)produce and transform facts and truths (ibid, pp. 111-113). I see my study as mainly corresponding with this latter understanding of critical thinking. Limberg et al find a

discursive shift in a broader understanding of literacy, evolving with digital development, in which people can be considered literate in various types of written and spoken documents and media. By bringing into view taken-for granted understandings and implicit assumptions, discourse analysis can help literacy education by providing different understanding of learning, information and technology (Limberg et al., 2012, pp. 113-115). I believe that this study can contribute to the understanding of today's information landscape and what literacies are required of the user to navigate in it.

In the anthology *Critical Theory for Library and Information Science: Exploring the Social from across the Disciplines*, Olsson (2010, pp. 67-68) emphasizes that research in library and information science has been criticized for mainly ignoring issues regarding social inequity and power relations. He argues that the dominant theoretical approaches within this field of research has been said to focus on the individual information seeker and not to provide any basis for theorizing on how social inequity and power relations can influence information behaviors and practices. However, the research that I have discussed here illustrates that there clearly are exceptions to this within LIS. Olsson stresses the relevance of Foucault's focus on discourse and power as fruitful for LIS research and acknowledges a growing interest for discourse analytic approaches (Olsson, 2010, pp. 63-64). He points out that by acknowledging discourse, the individual information seeker, or other concerns for LIS research, cannot be seen as isolated from the discursive context. He also emphasizes the relationship between knowledge and power – a salient feature of Foucault's work - as potentially important for LIS research (ibid, pp. 67-68).

2.2 Privacy in LIS

The view and challenges of patron privacy in the digital era has been discussed by Zimmer in "Assessing the Treatment of Patron Privacy in Library 2.0 Literature" (2013). He argues that as libraries embrace web 2.0 technologies, they are facing new dilemmas that challenge traditional ethics of librarianships regarding the protection of their patrons' privacy.

> In the library, users' intellectual activities are protected by decades of established norms and practices intended to preserve patron privacy and confidentiality […] Library 2.0 threatens to disrupt these norms. In order to take full advantage of Web 2.0 platforms and technologies to deliver Library 2.0 services, libraries will need to capture and retain personal information from their patrons.
>
> *Zimmer, 2013, pp. 30-31.*

Zimmer analyzes how privacy concerns are articulated within the professional discourse of library 2.0 by determining if, and how, patron privacy are introduced and discussed by librarians and information professionals within trade publications. Out of 677 articles discussing Library 2.0, Zimmer find that 36 articles (5.8 percent) relevantly mention patron privacy, only 11 in more than merely passing mention. However, over half of the 36 articles indicate a high or moderate level of concern about these issues but only 14 of them provide means to address privacy issues. Zimmer (2013, p. 36) writes that these results suggest while the potential for Library 2.0 to provide new content and services is accentuated within the professional

discourse, the discussion on how these adoptions and implementation of these tools might effect patron privacy is minimal.

> Consequently, as the interest in, and adoption of, Library 2.0 services increase, librarians and related information practitioners seeking information regarding these new technologies in professional publications will not likely be confronted with the possible privacy concerns, nor learn of any strategies to deal with them.
>
> *Zimmer, 2013, p. 36.*

This is highly connected with a study conducted by Burkell and Carey, laid out in "How well libraries Personal Information and the Public Library: Compliance with Fair Information Practice Principles" (2011). They state that libraries today collect and store many types of personal data and that this raises privacy risks for patrons. By examining if public libraries in Ontario have privacy notices and if they agree with recommended regulations, Burkell and Carey find that a majority of the 76 public libraries included in the sample fail to provide their patrons with notice as required by regulatory framework, and most of the libraries that attempt to do so in an ineffective manner. Fewer than half of the libraries offered any form of notice on how they collect and use the patrons' personal information.

> Privacy notice in the form of a privacy policy or notice required by regulation is not a panacea for privacy concerns. It is, however, a step in the right direction. By providing comprehensive notice regarding the collection and use of personal information, libraries allow their patrons to make informed decisions on the release of their personal information.
>
> *Burkell & Carey, 2011, p. 14.*

Online privacy and privacy regulation as information practice has, within the LIS-field, been investigated by Fred Stutzman. By deploying stages from Communication Privacy Management theory, CPM, Stutzman and Kramer-Duffield studies privacy as information practice in "Friends only: Examining a Privacy-Enhancing Behavior in Facebook" (2010). They examine the friends-only Facebook page setting as privacy-enhancing practice among undergraduate students and find that having one's page set to friends-only is associated with increasing levels of interpersonal privacy management practice and "expectancy violations", i.e. when the individual's expected audience is not the intended audience. I will also deploy CPM theory within this study as a way to analyze mechanisms of privacy.

How privacy behavior online has developed over time is elaborated by Stutzman, Gross and Acquisti in "Silent Listeners: The Evolution of Privacy and Disclosure on Facebook" (2012). By examining how the privacy and disclosure behaviors of over five thousand Facebook users have changed over time, from 2005 to 2011, three contrasting trends are identified. The users increased their privacy-seeking behavior, meaning they decreased the amount of personal data that they shared publicly with profiles they weren't friends with, or in other ways connected with. Stuzman, Gross and Acquisti explain this with for example: access to simpler privacy settings, growing expertise with settings and increased awareness of online privacy risks. By the end of the study, in 2009, a contrasting trend appeared due to changes in Facebook's privacy policy and interface settings, which produced greater non-intentional public disclosures by the users. While the authors explain the first trend as

user-driven due to the members' efforts to protect their privacy, the second trend, produced by Facebook, inverted the former trend. Lastly, the authors find that over time, the scope and amount of personal information disclosed "privately" to friends' profiles has largely increased. This also means that the personal information comes in the hands of what the authors call *silent listeners*; Facebook itself, third-party applications and (indirectly) advertisers. Stutzman, Gross and Acquisti (2012, p. 31) suggest that these findings: "highlight the challenges users of social network sites face when trying to manage online privacy, and the power of providers of social media services to affect individuals' disclosure and privacy behavior through interfaces and default settings."

The effects of privacy policy consumption and privacy settings on privacy attitudes and behaviors of disclosure are furthered developed by Stutzman, Capra and Thompson in "Factors mediating disclosure in social network sites" (2010). By a survey answered by 122 Facebook-users, they find that privacy attitudes do not affect disclosure practice, while increased privacy customization (if the user has customized her/his privacy settings) has a positive impact on disclosures on Facebook. Increased privacy policy consumption, on the other hand, has a negative impact on disclosures on Facebook. The authors stress that this calls for understandable privacy policies and usable privacy controls, and suggest that: "simplifying privacy policies and their presentation […] is an important part of helping users to feel more confident that they understand the range and implications of their disclosures" (Stutzman et al., 2010, p. 597). Their results also call for a deeper analysis of the discourse of the policies, as they evidently affect privacy practice.

Another example of privacy research within LIS, which takes the development of the last decade into account, is "Protecting Private Information: Current Attitudes Concerning Privacy Policies" by Williams, Agarwal and Wigand (2014). The study investigates what the current attitudes towards privacy policies are and if those attitudes have changes in the last ten years. By conducting an online survey the authors find that only 7 percent of the respondents always, or almost always, read privacy policies, while 70 percent never, or almost never, read them, mainly because they consider the policies too long or too complex. Among a third of the respondents thought that privacy policies were valuable or extremely valuable, while about the same number of respondents thought that privacy policies are worthless. 71 percent answered that they have decided not to provide personal information after reading a privacy policy. The study by Williams, Agarwal and Wigand also shows that there are large misconceptions and confusion about the purpose of privacy policies, since a large number (72 %) believed that, or did not know if, the existence of a privacy policy meant that their information would not be sold. By comparing their results with research from 2005 the authors find that the attitudes have not changed, and that no real change in privacy awareness has occurred. Williams, Agarwal and Wigand (2014, p. 6) argue that in the way the privacy policies are written; "they serve only to protect organizations from sharing or selling consumers' private information to other organizations." The authors further suggests that "organizations should strive to make the specifics of how they handle private information clear and obvious to their users, without (a) hiding behind words such as share when in reality they mean sell" (Williams et al., 2014, p. 7).

2.3 Power, Privacy, and Policies in other Fields

> We need to make sure that we are not creating a "digital literacy divide", that we are not failing vulnerable populations by allowing them to be exploited. For this reason we examine the state of literacy online and whether privacy policies live up to standard of being understandable to everyone.
>
> *Jensen & Potts, 2003, p. 2.*

Privacy policies has been investigated from a usability perspective by Carlos Jensen and Colin Potts, in "Privacy Policies Examined: Fair Warning or Fair Game?" (2003). Although the report can be considered old in the fast-changing Internet development it offers results about the readability of privacy policies that still lives on. Jensen and Potts states that it is important to examine the literacy online, and the privacy policies, since they are the main source of information that the users can base their decisions on. After the policy has been posted and made publicly available the user is solely responsible for her own protection, which Jensen and Potts (2003, p. 1) states makes the practice of privacy policies compelling to businesses since it do not require much of them. The authors also emphasizes the non-negotiability of the policies: "The user is presented with a set of terms and conditions, and has no leverage, or voice to negotiate new terms" (Jensen & Potts, 2003, p. 1). By using standardized readability measures, they analyze the privacy policies of 22 medical and healthcare related websites. They find that the policies on average require reading skills equivalent with "some college education" and that about a third of the adult Internet users over 25 years only posses the skills to understand one of the policies. Jensen and Potts come to the conclusion that nearly half of the Internet users in the US do not have reading skills to make sense of the average privacy policy, which means that the policies fail to meet their purpose. This relates to information literacy, as well as discourses of it, as previously discussed by Limberg et al (2012):

> This indicates the presence of a significant digital literacy divide, at least in the context of privacy protection. As the Internet has reached a larger, more diverse audience, more and more people are left behind, and left vulnerable to abuses of their personal information.
>
> *Jensen & Potts, 2003, p. 5.*

The time that it requires of a user to read policies are also significant and the user still cannot tell whether the companies abide to their policies, since they can not know what happens behind the scene. Jensen and Potts also comment on the practice of assuming that everyone using the service has given consent to the privacy policy, for this to be fair the policies would have to be accessible through "safe areas" of the website, areas free from all kinds of collection:

> Without such provision, the simple act of consulting the policy (requiring the user to at least access the policy page and the site's front page) means that the users have already given their consent, a "Catch-22" situation. This practice violates the very essence of the concept of fair warning as well as consent.
>
> *Jensen & Potts, 2003, p. 6.*

They also argue that while the companies only have to formulate one single policy, the user is expected to review every policy of every company they happen to interact with.

> When it comes to the issue of availability, most websites featuring a privacy policy provide (at least) a link to it from their main page. Though not always prominently featured, sites which do have a privacy policy usually make it available to users, at least those users who know to look for it. This is an important issue: A privacy policy may not be sought out or consulted unless a user suspects information about them is being collected.
>
> *Jensen & Potts, 2003, p. 1.*

Policies may well have become user-friendlier since their study was conducted but the premises remain the same. Jensen and Potts (2003, p. 8) writes that users are unlikely to follow links to privacy policies and review what they say, and may make the wrong assumption of believing that the fact that a site has a privacy policy should in some way mean that they protect users' privacy.

Helen Nissenbaum and Solon Barocas discuss privacy in the context of big data in "Big Data's End Run around Anonymity and Consent" (2014). They state that it is understandable that anonymity and consent are attractive tools to protect privacy, anonymity since it means that data no longer can be related to identifiable subjects and consent since it matches the dominant conception of privacy as being able to control information about oneself. However, Nissenbaum and Barocas consider that these tools seldom work in practice: "anonymity and consent have proven elusive, as time and again critics have revealed fundamental problems in implementing both" (2014, p. 45). They argue that values that anonymity protects are undermined by common applications of big data; although individuals are not directly identifiable, they may still be able to reach through details of attributes and activates in records. Regarding consent they argue that it is inefficient as a matter of individual choice and that it is absurd to believe "that notice and consent can fully specify the terms of interaction between data collector and data subject" (ibid, p. 45). Nissenbaum and Barocas argue that the need to protect privacy mainly is perceived as finding ways to support notice and choice, which has been mended by offering individuals unilateral terms-of-service contracts in the form of privacy policies. These has remained the core of privacy protection despite evidence that few read and understand them, according to Nissenbaum and Barocas (2014, p. 57). They state that regulatory agencies have demanded improvement in the ways privacy policies are expressed and communicated, but that this stands in contrast to what Nissenbaum and Barocas call the *transparency paradox*. Nissenbaum and Barocas stress that while there is a demand for plain language and simple-to-understand privacy policies, this cannot be combined with the demand for transparency, and "even when people understand the text of plain-language notices, they still will not – indeed cannot – be informed in ways relevant to their decision to consent" (ibid, p. 59).

Several aspects of privacy on social media services are discussed in the anthology *Privacy online: Perspectives on Privacy and Self-Disclosure in the Social Web* (Trepte & Reinecke, 2011). In one of the chapters Debatin (2011) discusses the ethical basis for privacy and self-restraint in social networking online. He examines the definition of privacy and highlights its moral and ethical value, defined by the

United Nations as a human right. Privacy protection is, according to Debatin (2011, p. 48), becoming vital in the informational dimension because of the fast advances in information technology and its processing and storing capacity. He states that privacy can be protected by legal regulation, ethical self-regulation (with informational norms) or privacy-enhancing technology. However, Debatin argues that privacy-enhancing technologies have questionable reliability and trustworthiness, that they are not sufficient and only creates a false security. Although Debatin argue that user privacy should rightfully be protected, he stresses that privacy protection in social media seems to be an oxymoron:

> After all, the main purpose of participating in social networks is the exchange of information, most of it highly personal, and the maintenance and expansion of one's social relationships.
>
> *Debatin, 2011, p. 54.*

Debatin argues that privacy risks while using social networking sites are privacy risks at the *horizontal axis*; social interactions that the user are more or less aware of, and at *vertical axis*; systematic collection and use of data, that mostly remain invisible to the user. Debatin calls the risks at the horizontal axis the tip of the iceberg, while the risks at the vertical axis proposes the greater problem. Most of the users are, according to Debatin (2011, pp. 55-56) aware of some risks, but do not act accordingly and are often unaware of privacy policies, as well as satisfied with the mere idea of privacy control without much real content. Debatin (2011) promotes the implementation of *privacy literacy*, by which he refers to that users should develop knowledge to be able to see through the "technological veil" and make educated choices. He states that social media users need to develop informed concerns about their privacy, inform themselves about potential negative effects social media can have on their privacy and acquire skills to prevent or mitigate the negative consequences.

> And finally, ethicists, educators, system developers, and service providers are also responsible for creating an environment that fosters privacy literacy among the users of social media and in society as a whole.
>
> *Debatin, 2011, p. 58.*

Differences in European and American perspectives on privacy are investigated by Nina and Boers in "Disliking the like: User policy-change and perception of the internet as a democratic medium" (2013). They argue that privacy is seen as a human right in European policy whereas the US approach is much more consumer driven, as they state: "Consumer confidence and trust are the primary focus of privacy policy." (Nina & Boers, 2013, p. 322). Also, they argue that US lack a unified approach. However, they write that the user's right in European policy is mostly made of the "right to be forgotten", which they discuss appears as a symbolic right since it today is technically impossible to completely delete the private information. Nina and Boers states that one could question whether citizens needs to be protected from privacy violation, since the citizens actually are "consumers using commercially provided services with policies to which they have agreed" (Nina & Boers, 2013, p. 320). However, Nina and Boers discuss whether most users have the proper level of media and information literacy skills to manage their own privacy online. They also write that many users depend on the provided service at the same time as they may have

objections to the service's core policies. Nina and Boers (2013) highlights how users may take action to change privacy and data policies and mention privacy policy changes for LinkedIn that caused an uproar among its users and made the company recant the change. This suggests that users could have a greater say over the policy content, but Nina and Boers state that in the attempts of user-governance that have been made users show little interest.

Similar conclusions are made by Bechmann, in "Non-informed Consent Cultures: Privacy Policies and App Contracts on Facebook" (2014), who discusses non-informed consent cultures with a case study of consent and privacy concerns among a group of high school students. She emphasizes that statistics from EU shows that most users of social networking sites do not read the privacy policies and in her study none of the students had read them, which she argues collides with the legal interpretation of informed consent that allow companies to circumvent personal data handling regulations.

> There is a long way from the idea of 'informed consent' as an isolated agreement that has been accepted once or is being accepted on a regular basis to the practice of skipping information in order to get to the service, relying on the consensus among group members, or to simple engage in a gift economy where the Facebook individual has to deliver data in return for socialization.
>
> *Bechmann, 2014, p. 35.*

The high school student in Bechmann's study that had most privacy concerns described an ambivalent situation where the student need to accept permission that the student found inappropriate, in order to be part of the online socialization taking place in the group of friends. Bechmann (2014) also discusses transparency, arguing that there is a problem in the need to know more about the use of data and the information overload it creates for the user.

Power in terms of privacy and discourse on Facebook has been investigated by Buchanan in *Privacy and Power in Social Space: Facebook* (2011). She examines participation, Facebook as platform and the company that operates it, as well as developers of applications and political parties' election campaigns on Facebook. She states that users are pressured to upload content, personal information, thoughts, preferences and relations, and that Facebook depend on this user-generated data. The users produce and consume content and this is also what generates income for Facebook, since the company collects personal information and passes it on to external companies such as advertisers. She emphasize that Facebook uses the term "share" while it more correctly is selling the information that the company is doing (ibid, pp. 273-274). Further, Buchanan (2011, p. 277) stresses that power is manifested in the surveillance of users' content, which together with all user activity is recorded and analyzed.

> Users' power to control access to their personal information has been greatly reduced by successive technical programs introduced by Facebook Inc. that enable external companies and organizations to determine to which information they want access, and to refuse access to their applications if users do not comply.
>
> *Buchanan, 2011, p. 280.*

She thus argues that this is the most important way in which power is manifested on Facebook, by access to personal information and data and the ability to deny full participation in the network if not agreed upon (2011, pp. 280-281).

2.4 Summary

The previous research presented in this chapter provides perspectives on power, discourse and privacy, within LIS and within other fields of research. But it also provides perspectives on media and information literacy, naming, and privacy in the development of big data, as well as power structures in social media and viewpoints on privacy policies and privacy settings. To form a basis for my analysis, in relation to previous research, I will in the following chapter elaborate my theoretical framework.

3. Theoretical Framework

> nobody who has an interest in modern society, and certainly nobody who has an interest in relationships of power in modern society, can afford to ignore language.
>
> *Fairclough, 2001, pp. 2-3.*

For my study I implement a theoretical framework on a general level based on discourse theory, which will frame the study and entails a theoretical viewpoint of the concept of power. For my investigation of the notion of privacy and to understand power mechanisms of privacy, theoretical approaches to the concept of privacy are also implemented.

3.1 Theoretical Perspectives on Discourse and Power

Different kinds of discourse analyses have been developed, as previously mentioned in the background chapter. One of them is the Critical Discourse Analysis, henceforth CDA, which has an approach that is based on social criticism and power criticism. That is why it constitutes my theoretical framework, since power structures expressed in language are the main research object of this study. According to Boréus (2013b, p.153) CDA, as well as the other kinds of discourse analyses, considers that how we talk and write about phenomena affects other social practices. But in comparison to more foucauldian-oriented analyses, CDA further emphasize how these other, non-linguistic (non-textual), social practices also affect language use. CDA: "incline to critical realism rather than post-structuralism and focus analysis on relations between discursive and material elements of social life rather than just discourse" (Fairclough, 2013, p. 177).

3.1.1 Theory of Critical Discourse Analysis

CDA defines discourses as mainly affecting how we perceive reality and understand society but also the shaping of identities and relations between groups (Boréus 2013b, p. 153). One of the main contributors to the development of this analysis is Norman Fairclough, who argues that regularities and expressions of language use, which constitutes discourses, are manifested in semiotics, as texts. In Fairclough's own words: "social relations, power, institutions and cultural practices are in part semiotic, they internalize semiosis without being reducible to it" (2013, p. 179). This thesis is based on the notion that power structures can be, and are, expressed and shaped by language. In *Language and Power* Fairclough (2001) lays out a theory of how power functions through language and that there is an internal and dialectical relationship between language and society. By this Fairclough argues that when people use language by writing, talking, listening or reading they do so in ways that are socially

conditioned and have social effects. Furthermore, language takes place in social contexts and is not just an expression of social processes and practices - it is part of it. For example, Fairclough acknowledges that political struggle occurs in language and over language. To specify this relation between language and society, Fairclough states: "whereas all linguistic phenomena are social, not all social phenomena are linguistic – though even those that are not just linguistic (economic production, for instance) typically have a substantial, and often underestimated, language element." ((Fairclough, 2001, p. 19). He states that sociolinguistic conventions incorporate power asymmetries and also arise out of, and contribute to, certain relations of power. The theoretical approach of Fairclough (2001, p. 2) emphasizes what is viewed as common-sense assumptions, what is implicit in sociolinguistic conventions and is manifestations of ideology. This, he argues, occurs as the conventions are used as means to legitimize existing social relations and power differences.

Power by consent, in terms of ideological workings of language, is done by practices that Fairclough (2001, p. 27) stresses can seemingly be universal or based on common sense. That means that they can be originating from the dominant class or bloc, but people draw upon these practices without thinking – the practices have become naturalized. When practices, and types of discourses, function in this way to sustain unequal power relations, they are functioning ideologically. Fairclough further defines ideological power as a complement to economic and political power and consisting of the power to: "project one's practices as universal and 'common sense'" (Fairclough 2001, p. 27). Power by consent is exercised by winning other's acceptance for one to possess and exercise power, and I believe that the privacy and data policies could be a manifestation and a part of this specific exercise of power.
Power by consent is according to Fairclough (2001, p. 30) increasing, and this makes it possible to implement increased practice of social control: "This is often a matter of integrating people into apparatuses of control which they come to feel themselves to be part of (e.g. as consumers […])." The key element in power by consent, and thereby through ideology, is discourse. Fairclough (2001, p. 73) states that ideological struggle mainly takes place in language and over language, since language is what's at stake as well as the site of the struggle.

Fairclough makes a distinction between power *in* discourse and power *behind* discourse. The first concerns: *powerful participants controlling and constraining the contributions of non-powerful participants* including constraints on what is said or done (content), the social relations people enter into in discourse or the subject positions (social roles) people can occupy (Fairclough, 2001, pp.38-39). Today a large amount of discourse involves participants who are separated in place and time, Fairclough (2001, p. 41) argues. The main difference of this versus face-to-face discourse is the one-sidedness, as with traditional media, where there is a sharp divide between producers and interpreters (ibid, p. 41). Whereas in face-to-face discourses the producer can shape the language to fit the person in front of them, the media discourse is made for mass audiences and therefore addresses an ideal subject. This is something that the policies which this thesis will investigate has to do as well, since they are shaped for a broad audience they cannot customize their language but can be expected to address an "ideal" image of a user.

Fairclough (2001, p. 43) labels the power relations of the media as a mediated sort, between the power-holders and the population. Because this mediated power is implicit, Fairclough considers it to be hidden power. The grammatical examples of how this works concerning mass media can also be used to study the power relations in the policies. One focus is *causality*: "who is represented as causing what to happen, who is represented as doing what to whom" (ibid, p. 43). This is linked to the form of *nominalization* when a process is expressed as an entity, a noun. Fairclough argues that one effect of using this is that aspects of the process – such as causality - are unspecified.

> The power being exercised here is the power to disguise power […] It is a form of the power to constrain *content*: to favour certain interpretations and 'wordings' of events, while excluding others.
>
> *Fairclough, 2001, p. 43.*

Power *behind* discourse is: "that the whole social order of discourse is put together and held together as a hidden effect of power" Fairclough (2001, p. 46). This is the power effect that imposes discourse types upon all those involved, discourse types that Fairclough (2001, p. 51) states do not belong to the institution or system in itself but to the power-holders in it and the ideology behind them. One aspect of power behind discourse is also the question of access: who has access to which discourses, and who has access to discursive positions of power: the power to impose and enforce constraints on access to the discourse (ibid, p. 52). Fairclough (2001, p. 53) also mentions literacy as one important factor that is unequally distributed and precondition to access discourse. Fairclough (2001, pp. 54-57) consider it to be a factor keeping access restricted by making high demands on participants that can be hard to meet. Discourse in formal settings can be difficult and depend on special knowledge and skill.

To conclude, Fairclough (2001, p. 61) argues that regarding power *in* discourse the discourse is the site of the power struggle, whereas regarding power *behind* struggle the control of the discourse itself is what is at stake. This means that the focus of struggle is the establishment or maintenance of one discourse's domination over other discourses in a social domain. The dominating discourse thereby establishes or maintains its particular ideological assumptions as commonsensical (ibid, p. 75). One aspect of the commonsensical discussed by Fairclough (2001, pp. 77-84) is the meaning of words. There is a belief of words as having fixed meanings but Fairclough argues that meanings can vary between social dialects and ideologies. He states that a meaning of a word is not isolated and independent but part of a meaning system, the relationship of similarity, contrast, overlap and inclusion as well as a words relation to other words. When the meaning of a word or concept becomes fixed, it can be seen as an effect of power by the ideological effect of naturalization (Fairclough 2001, p. 78, 89). Further, Fairclough (2001, p. 89) states that a dominant discourse itself is undergoing a process of naturalization, in which it appear to lose its connection to any particular ideology and interests, and begin to be considered as common sense: "The naturalization of the meanings of words is an effective way of constraining the contents of discourse and, in the long term, knowledge and beliefs" (ibid, p. 87).

3.1.2 Interface as Discourse

To fully understand the power structures expressed through the policies the context that surrounds them also needs to be taken into account. Therefore, I combine my discourse analysis of the content in the policies with a discursive interface analysis. The purpose of this is to investigate in what web context the policies can be found. I see this as an extended and important part of what the companies and services mediate through their policies. Stanfill describes this analysis in "The interface as discourse: The production of norms through web design" (2014) and argues that: "This lens allows examining the assumptions built into interfaces as the normative or 'correct' or path of least resistance" (Stanfill, 2014, p. 2). By starting in Foucault's notion of power as productive, Stanfill states: "A productive power framework operates from the premise that making something more possible, normative, or 'common sense' is a form of constraint encouraging that outcome" (ibid, p. 2). Further meaning:

> The interface makes a normative claim; it is not an omnipotent system. Not every site visitor responds in the same way, but to understand the norms sites produce, analysis must consider which responses become the path of least resistance and how.
>
> *Stanfill, 2014, p. 3.*

Stanfill argues that by analyzing interfaces we can find how certain uses are easier while others are harder, exposing norms of use: "discursive interface analysis takes a critical perspective attentive to unequal power between industry and site visitors." (Stanfill 2014, p. 4). Aspects of a interface discourse analysis that can be useful to my analysis is the ones concerning ease to distinguish features, to take into account how these policies can be reached. These are naming and labeling, as labels on a button or name of a menu, leading to the policies and page placement, as the placements of the button, link or menu leading to the policy in the services' online interfaces (ibid, pp. 5-6). Stanfill (2014, p. 6) argues that something appearing on the top or left is more visible than on the lower or on the right (by left-to-right reading standard). What cannot be seen without scrolling down is easily overlooked and higher placement indicates weight and visibility. Making something stand out, or not, by design choices assumes the valuation of that information and the design choices "both reflect and reinforce assumptions and valuations" (ibid, p. 6).

The theoretical concepts elaborated by Stanfill provide me with analytical approaches to the interfaces in which the users can reach the privacy and data policies. By this critical theoretical perspective on interfaces I can deepen my analysis of the power structures of the privacy and data policies and their contexts.

3.2 Theoretical Perspectives on Privacy

How privacy can be perceived and the mechanisms of privacy boundaries are the main concerns of Communication Privacy Management theory, henceforth CPM, established by Sandra Petronio in *Boundaries of Privacy – Dialectics of Disclosure* (2002). CPM can be applied to many different contexts and is today especially used regarding issues of information technology and social media (Vasalou, Joinson and Houghton 2014, p.3). Margulis (2011) argues that CPM is the most valuable theory

for understanding interpersonal computer-mediated communication and especially suited for studies concerning social networking. Stutzman and Kramer-Duffield (2010) deploy the CPM theory on a user-to-user level while I intend to use it to analyze the discursive notion of privacy, corresponded by the companies addressing the users, as expressed in the policies. CPM considers privacy and disclosure to be inseparable features of a unified dialectical process, which Altman (2002, p. xv) argues is a fairly recent idea stemming from the 1970s. Altman emphasizes that this balance is the core in Petronio's theory:

> How do we strike a balance between the incredible positive opportunities to reach out to others made possible by modern technology, versus the dangers of losing the ability to control and regulate what others may know or have access to about us?
>
> *Altman, 2002, p. xiv*

Petronio argues that the decision to reveal personal information never is straightforward:

> We are constantly in a balancing act. We try to weigh the demands of the situation with our needs and those of others around us. Privacy has importance for us because it lets us feel separate from others. It gives us a sense that we are the rightful owners of information about us.
>
> *Petronio, 2002, p. 1*

Central to CPM is the viewpoint that individuals believe they "have a right to own and regulate access to their private information" (Petronio, 2002, p. 2). In CPM *boundaries* are used as metaphor to illustrate how borders of ownership control the flow of information to others, according to Petronio (2002, p. 3). By CPM this regulation process is considered communicative, and communication as the core of private disclosure because of the theory's focus on the process of granting or denying access to private information. CPM "intersects the individual with the collective to gain a broader view of a specific communication phenomenon where people manage private information" (ibid, p. 23). In this sense I regard the privacy and data policies as communicative, a part of the communication between company and user, and at the same time the objects they deal with are also platforms of communication – social media. Petronio (2002, p. 3) states that by the privacy management system CPM theory identifies ways that privacy boundaries are coordinated between agents, in my case between a service/company and user, which offers ways for me to deepen the analysis of how the privacy boundaries are formulated in the policies (regarding my research questions) and what it suggests regarding the notion of privacy. Petronio states that in CPM theory "privacy is defined as the feeling that one has the right to own private information, either personally or collectively: consequently, boundaries mark ownership lines for individuals" (ibid, p. 6).

CPM is based on five theoretical suppositions, laid out by Petronio (2002, pp. 3-13):

- The first is the focal point of private information as the content of disclosure
- Second, the boundary metaphor illustrates the distinction between private information and public relationships
- Third, one core aspect is control because "people believe that private information is *owned* or *co-owned* with others: thus they desire control over

the boundaries" (Petronio, 2002, p. 3). Also, control is important to diminish the vulnerability that may follow revealing or concealing private information

- Fourth, to understand how boundaries are regulated the theory uses a rule-based management system
- Fifth, privacy and disclosure are treated as aspects of a dialectical process

Within the framework of CPM theory, people consider private information as something they own and desire control of, as well as something they either can reveal or conceal (Petronio, 2002, p. 9). The theory suggests that the rule management for privacy boundaries is something that every individual engages in, either consciously with outlined principles or unconsciously as they go along making decisions regarding their private information. The ownership boundaries for private information may be clear or they may be ambiguous, and the boundaries function to distinguish ownership of the private information and control who can know about private matters. According to Petronio, we consider it a violation of privacy when someone else is trying to control information that we perceive as ours. By controlling private information we try to protect ourselves against exposure. Information that we have shared become co-owned and as we are being told private information, CPM sees it as we agree to a contract of responsibilities and that the choice for the ones involved to disclose or conceal information depend on a risk-benefit ratio (ibid, pp. 6-10). Petronio (2002, pp. 10-12) suggests that the rule-based management system gives us a structure for understanding of how private information is handled. Furthermore she suggests that the system works on both personal and collective levels. According to Petronio (2002, pp. 4-5), once disclosure takes place people become involved in collective boundary management, since the co-ownership of that private information calls for a coordination of different privacy rule foundations: "people within the boundaries must coordinate with others so that the rules are known and used according to agreed-upon ways" (ibid, p. 19). The rules concern boundary linkage, ownership and permeability. Petronio (2002, pp. 30-31, 77) writes that identification of ownership and co-ownership can be unclear, which may lead to conflict over ownership issues. She argue that since the one disclosing is the original owner of the information that person often feels a right to be able to determine which rules that should be used regarding third-party disclosure.

The collective coordination pattern, acknowledged by CPM, which correspond with the relationships in this study is that of *inclusive boundary coordination*, in which power is a key component since it is used to describe when one person gives up privacy control to another person (or as in this study: a company), which leads to increased vulnerability. Petronio (2002, pp. 127-131) suggest that this can be defined by three different boundary linkages:

- *Coercive linkages* when the person is forced to reveal private information to another person
- *Role linkages* when individuals (or companies) hold positions that dictate who gets access to information
- *Susceptibility linkage* when persons do not monitor themselves and end up in a situation where they disclose more information than their recipients do.

Individuals who find themselves in the last linkage, susceptibility linkage, can potentially be in a vulnerable situation where the recipient has more power to control

the discloser's information than the discloser has: "it leads to power incongruity where the confidant […] is privy to a great deal of personal information that has the potential to become a power source" (Petronio, 2002, p. 130).

Within inclusive boundary coordination the ownership can be defined in different ways, such as *benevolent* ownership where the confidant sees it as their responsibility to take the disclosers wishes for third-party revelations into account, or *manipulative* ownership where the confidant sees the co-ownership as a way to completely dominate how the information is managed (Petronio, 2002, pp. 130-131). The confidant has more influence over the access to the discloser's information than the discloser has over access to her or his own information or to the confidant's private information. Petronio (2002, pp. 131-132) writes that this power discrepancy results in the private information that the discloser has revealed being under the command of the confidant, who can make decisions of revealing or protecting the information. This puts the confidant in a power position where the discloser must persuade the confidant if she or he wishes to change the rules for concealing or revealing.

> If we think the information belongs to us, we perceive we have the right to regulate it according to our own rules of revealing and concealing. Turbulence arise when others see the same information as collectively owned and managed according to mutually established rights.
>
> *Petronio, 2002, p. 190.*

When the coordination of people's privacy boundary rules does not work satisfactory there is a risk of *boundary turbulence*. Petronio (2002, p. 177) argues that there are times when this misstep is apparent to those involves and measures are taken to make the coordination better, but that it as well are times when the difference in rules usage and the miscommunication is not clear until it causes a conflict. Petronio (2002, p. 21) argues that boundary turbulence may be the result of such as a misunderstanding of the rules, when co-owners are believing in differing rules, applying boundaries from other boundaries or ignoring the collective rules. She writes that when boundaries are fuzzy the confidant or to-be co-owner does not understand the rights and responsibilities for the management of the information (ibid, pp. 21, 31, 177-180). She argues that fuzzy boundaries occur: "when people are ambiguous about who owns or co-owns the private information, changing the rights to determine rules" (ibid, p. 190).

> If the discloser only hints at a rule for how to access or protect the private information, the confidant may not really understand the way he or she should treat privacy management. This uncertainty may result in misunderstandings and hurt feelings when the rule is not applied in the way the disclosure envisioned.
>
> *Petronio, 2002, p. 78.*

In my case, since I investigate the policies concerning the managing of information about others: the hints at rules may instead be done by the to-be co-owner of information – the confidant in form of the company – and the one who may or may not understand how the privacy management will be treated is the user in the form of discloser. Petronio also refers to inappropriate claims as a factor that explains why some breach confidences and can't understand why others complain about it.

> In their minds, the shared private information becomes solely theirs to do with as they wish. Turbulence erupts because this assumption is rarely acceptable to other co-owners. Ignoring the mutuality of the information and acting in ways that contradict presumptions of dual ownership are problematic for boundary coordination.
>
> *Petronio, 2002, p. 180.*

She argues that the ownership can be questioned after private information has been disclosed; who is in the most right to call ownership and by that has the right to regulate the flow of the private information to others? Petronio suggests that boundary turbulence may also the cause of rule mistakes when people make errors in judgments. One is when people apply privacy management rules that do not agree with other member's perspective. They misunderstand rule expectations and assume that when information has been disclosed confidants will not tell others or that confidant will follow what the discloser believes is established rules. The discloser feels a sense of security, which can be deceiving. Another error in judgment may be when people do not pay attention to rule development or do not initially understand the privacy rules: "people can give boundary access irresponsibility when they misconstrue the collective rules for revealing or concealing" (Petronio, 2002, p. 185).

The key aspect of CPM theory is that disclosing and concealing private information is a dialectical process. The dialectical relationship of disclosure and privacy is correlating with openness – closeness and autonomy – connectedness. It considers the dialectical tension of "the needs of being both private through concealing and public through revealing" (Petronio, 2002, p. 12). Margulis explains that the dialectical aspect in CPM is important since:

> we continuously adapt our level of privacy and disclosure to internal and external states because we simultaneously need to be open and social as well as private and preserve our autonomy.
>
> *Margulis, 2011, p. 12.*

3.3 Summary

In conclusion, the theoretical perspectives of discourse and privacy complement each other and offer me a set of theoretical tools and viewpoints by which the privacy and data policies and their contexts can be analyzed. CDA provides a theoretical understanding of power and language as well as analytical tools to approach the texts; such as power in and behind discourse, grammatical functions, and explicit and implicit statements. The discursive understanding of interfaces offers viewpoints of how online design and structural choices indicate power relations and normative values, and analytical tools by focusing on naming, labeling, placement, and visual design of links and menus by which the policies can be reached. Lastly, CPM theory offers ways to understand the notion and mechanisms of privacy, by concepts such as boundary coordination, boundary linkages, (co-)ownership and boundary turbulence. In the following chapter I describe how CDA is methodologically deployed in this study.

4. Method

This chapter aims to clarify how this study is carried out. First I discuss how critical discourse analysis is conducted methodologically. This is followed by a description of what empirical material this study examines and its limitations. To specify the analytical operationalization, this chapter ends with a description of my analytical approach.

4.1 Critical Discourse Analysis in Practice

The core part of my investigation consists of an analysis of texts, in the form of privacy and data policies of social media companies. This is accomplished by using discourse analysis and is mainly motivated by the nature of what I investigate, which is power structures. Bergström and Boréus (2008, pp. 306, 328) stress that discourse analysis is scientifically oriented towards questions concerning power. The topic of this thesis can be investigated in other way apart from analyzing qualitative texts. One could for example make quantitative as well as qualitative surveys asking people of their perception of privacy and ownership of user-generated data, but due to my estimation this has in many ways already been done. I also believe that critical examinations of official texts are important to reveal power relations: how the texts are elaborated, retained and the language that is used as part of this. Texts do affect society and the conceptions and values of members of the society, as well as relations between people and between groups. Texts also affect the groups themselves, who are perceived as belonging to the group and texts shape and sustain identities (see e.g. Boréus 2013a, pp. 131-132). By such an interpretation, Boréus (2013a, pp. 131-132) stresses that texts can be studied as expressions of reign perceptions and relations.

The objects of critical discourse analysis are, according to Fairclough (2013, pp. 178-179) simultaneously material and semiotic and the dialectical relations between the semiotic and material is important to emphasize.

> CDA brings the critical tradition in social analysis into language studies, and contributes to critical social analysis a particular focus on discourse, and on relations between discourse and other social elements (power, ideologies, institutions, social identities etc.)
>
> *Fairclough, 2013, p. 178.*

By examining language in terms of discourse and social practice, one is not only setting out to analyze texts or mere processes of production and interpretation. The main focus is the relationship between texts, processes, and their social conditions, by Fairclough's (2001, p. 21) words: "both the immediate conditions of the social context and the more remote conditions of institutional and social structures". By

CDA the investigation can be made on three different levels; text, discourse and social practice. At the textual level, any analytical tools can be used to investigate for example what the text mediates explicit and what it mediates implicit; what is taken for granted. At the discursive level, one can study how texts and discourses are influencing each other and practices regarding how texts are produced and consumed in relation to the discourses they manifests (Boréus, 2013b, p. 154). The last step of the CDA is the level of social practice. Boréus (2013b, p. 154) defines this as relating the semiotic/linguistic practice (language use) to the social context. She claims that it can be hard to distinguish where an analysis at the textual level ends and the discursive level begins as well as to distinguish which processes belongs to the discursive level and which belongs to the level of social practice.

This study is made on each single privacy/data policy at the textual level. The discursive level consists of how the policies relate to each other and the interfaces in which they can be reached. How the discourses found in the policies and interfaces can be understood in relation to socio-political context, constitutes the level of social practice.

4. 2 Samples and Limitations

I have chosen to make my analysis on the privacy and data policies of different online services and applications that are rather commonly used by Swedish citizens and to which the user-generated data constitutes a large part of their product/content. I have decided which services by the help of for example Findahl (2014) and online sources (DeMers, 2015; Lenhart, 2015). I will analyze the privacy and data policies and interfaces of the following companies and services: Facebook, Google/Youtube, Twitter, Instagram, LinkedIn, Pinterest, Snapchat, Reddit, Tumblr and Ello.

The selection is also based on the intention to reach some breadth and variety in terms of the intended use of the services and their business profiles. Reddit stands out since the user do not has to provide any personal information to create an account, not even an e-mail address, and it is written in open source code (Reddit, 2015f). Ello is a new service and is not chosen due to common use but because it, so far, is a social networking service that do not sell advertisement and do not gather, collect or sell user data (Ello, 2015d). The reason I write Google slash Youtube is because Google owns Youtube, and by clinking on Privacy at Youtube, you get directed to the Privacy policy of Google (Youtube, 2015a). The choice of samples is also made with consideration of time and language skills, which is why for example Weibo is not part of the sample.

Overall the study is conducted on 10 policies. Besides from these texts I will also analyze the online context of the policies in 11 interfaces in web browser prior to login, 10 interfaces in web browser when logged into the service (since one can not log into Snapchat through web browser) and 10 interfaces in mobile phone applications (since Ello did not have a mobile phone application by the time of the study), as shown in Table 1. The policies are available online on the services websites and are thereby gathered by browsing their websites and saving the policies. The collection is made in the beginning of March 2015 and changes made in the policies, or to the interfaces where they can be found, after that point are not considered in this

study. In most cases the policies are called Privacy policies but it also occurs under the title Data policy, used by Facebook. Since though 9 out of 10 policies are titled Privacy policy this is mainly the name I use henceforth to describe general features of the policies.

Table 1: Empirical Material

Social Media Services	Privacy & Data Policies	Websites Pre Login	Websites Post Login	Mobile Phone Applications
1. Facebook	√	√	√	√
2. Instagram	√	√	√	√
3. Twitter	√	√	√	√
4. Tumblr	√	√	√	√
5. Pinterest	√	√	√	√
6. Snapchat	√	√		√
7. Ello	√	√	√	
8. Reddit	√	√	√	√
9. Linkedin	√	√	√	√
10. Youtube	√(Google's)	√	√	√
11. Google	√	√	√	√

4.3 The Social Media Companies

Facebook is a social networking service founded in 2004. It is a platform made for sharing content: text, images, videos, invites people to events etc. Its users are individuals, companies and organizations. Facebook states that its mission is to: "give people the power to share and make the world more open and connected" (Facebook, 2015e). According to the company it has over nine thousands employees and 890 million users that are daily active, 1.39 billion monthly active users. Since 2012, Facebook owns the service **Instagram** (launched in 2010), but the services do not share privacy policies. Instagram is a service made for sharing photos and short videos and is described as "a fun and quirky way to share your life with friends through a series of pictures" (Instagram, 2015d). According to the company the service has 300 million registered users.

Twitter is a micro blog, which allows its users to send short texts, attach photos, links etc., incorporated in 2007. According to the company it has 288 million monthly users and 3600 employees and claim its mission is to: "give everyone the power to create and share ideas and information instantly, without barriers" (Twitter, 2015d). **Tumblr** is founded in 2007 and is a blog platform service, on which the users can share photos, videos, links, text etc. According to the company it has 230 million blogs and 280 employees (Tumblr, 2015d). In 2013 Tumblr was sold to Yahoo Inc. (Tumblr, 2015c). **Pinterest** is a social bookmarking service founded in 2010. According to the company it has over 500 employees and is available in over 30 languages (Pinterest, 2015d). The service is described as: "a place to discover ideas for all your projects and interests, hand-picked by people like you" (Pinterest, 2015e). The service has not released any statistics of how many users the service has.

Snapchat were launched in 2011and is a service through which the users can send photos to each other, viewable for a restricted short time (Snapchat, 2015d). The company has not been revealed how many users the service has, but it is speculated to be from 100 million monthly active users to over 200 million users (Shontell, 2015).
Launched in 2014, **Ello** is a social networking site that profiles itself as being free from ads and not selling user data. It is a public benefit corporation. According to the company, there are "millions of people" using it, but no definite numbers is revealed (Ello, 2015d). **Reddit** is an open source community where the users can post stories; links, text etc. and vote content up or down. It was launched in 2005 and has nearly in mars 2015 it had 170 million unique visitors (Reddit, 2015e). The service is described as "a source for what's new and popular on the web. Users like you provide all of the content and decide, through voting, what's good and what's junk" (Reddit, 2015f).

Linkedin is a professional network launched in 2003, with the mission to: "connect the world's professionals to make them more productive and successful. When you join LinkedIn, you get access to people, jobs, news, updates, and insights that help you be great at what you do" (Linkedin, 2015d). According to the company it has nearly 350 million members and almost 7900 employees (Linkedin, 2015e).
Youtube was launched in 2005 and is a platform for sharing, watching and comment videos. It is owned by Google Inc. (Youtube, 2015b). The service claims it has over one billion users (Youtube, 2015c). **Google** was founded in 1998 and started out as a search engine (Google, 2015g). Now the company's services also includes, for example; e-mail client (gmail), web browser (chrome), maps (Google maps), research publications (Google Scholar), blog platform (Blogger) and the social community service Google+ (Google, 2015h).

4.4 Analytical Approach

The analytical approach of this thesis is based on qualitative deep analysis. By using discourse analysis I will scrutinize texts, in which discourse are manifested, to answer my research questions. In other words, the policies are seen as bearer of societal discourse. Through the analysis of usage of language in social context I will investigate discourse in terms of how it contributes to create and sustain power relations. The main study object is texts, but the contextualization of these texts is also of great importance for the analysis.

I find it important to also emphasize my role in the analytical process. When interpreting the sources I will inevitable carry with me all sorts of aware and unaware preconceptions. Like the case is with all research, I will accentuate what I find relevant in terms of my research aim and research questions, and decide not to include what I find irrelevant for this study. My focus in the analysis of the texts is, in accordance with my aim, on depictions relating to privacy, ownership, storing and use of user-generated data.

This study is made in three stages based on the analytical suggestions of Fairclough (2001, p. 91-93). Since it leaves some space for methodological choices and modifications I have pinned down an analytical approach that suits my aim and helps to answer my research questions. Foremost the three stages can be titled *description*, *interpretation* and *explanation*. By *description* Fairclough (2001, pp. 21-22) refers to

the analysis of the formal properties of the text. *Interpretation* can be seen as the first level of contextualization where the text is seen as "the product of a process of production, and as a resource in the process of interpretation" (Fairclough, 2001, p. 21). The final stage is *explanation,* which entails a second level of contextualization and focus on the "social determination of the processes of production and interpretation, and their social effects" (ibid, p. 22). The names of the three stages are also used as titles in the further presentation of my analysis.

First I will analyze my sources, the privacy and data policies, and distinguish expressions in the writing that relates to the research questions of this study, as previously mentioned. As described by Bergström and Boréus (2008, pp. 321-324) several different kinds of analytical tools for textual analysis can be deployed within CDA. I have here chosen to base my analytical questions on some of the ones suggested by Fairclough (2001, pp. 92-116) concerning vocabulary, grammar and textual structures, which suited my aim.

The questions that will make up my analytical guide while approaching the policies are the following:

1. What experiential values do words have?
    a. What classification schemes are drawn upon in the policies?
    b. Are there words which are ideologically contested?
    c. Is there rewording (one wording is replaced by another oppositional one) or overwording (high degree of words which are near synonyms)?
    d. What ideologically significant meaning relations (synonymy = same meaning, hyponymy = included in the same meaning, antonymy = meaning incompatibility) are there between words?
2. What relational values do words have?
    a. Are there euphemistic expressions (word replaced as a way of avoiding negative values)?
    b. Are there markedly formal or informal words?
3. What expressive values do words have?
4. What metaphors are used?
5. What experiential values do grammatical features have?
    a. Is agency unclear?
    b. Are nominalizations (a process converted to a noun) used?
    c. Are sentences active or passive?
6. What relational values do grammatical features have?
    a. Are there important features of relational modality (authority of the author in relation to others, expressed by for example verbs such as may, must, should, can)?
    b. Are the pronouns we and you used, and if so, how?
7. What expressive values do grammatical features have?
    a. Are there important features of expressive modality (statement of reality, are or are not)?
8. How are sentences linked together?
9. What larger-scale structures does the text have?

This consists partly of citations of questions stated by Fairclough, 2001, pp.92-93 and explanations given by Fairclough, 2001, pp. 94-116, modified by me.

Besides from the apparent meaning of text I will also look for what is taken for granted, in other words both the explicitly and the implicitly mediated. I will look for occurrence of words and concepts, in accordance to CDA exemplified in Bergström and Boréus (2008, pp. 341-342). Other linguistic aspects I will focus on are: *nominalization*: when verb or adjectives are replaced by nouns or when participants

in other ways are removed from the action described in a sentence, *passivization*: when a sentence is transformed in a way so that agents are removed and the focus becomes on the receiver of the process, and *modality*: in what way the sender relate to what is being said and can be considered to be tied to the message (Bergström & Boréus, 2008, p. 323; Boréus & Bergström, 2008, pp. 284-285). Finally, I will look for elements that are included and excluded from this discursive field.

The aspects and questions elaborated in my analytical guide are used to approach the empirical material. It is possible that certain questions regarding linguistic features may be found useful while others may not be present in any relevant form in the material and hence not further analyzed and discussed in the thesis.

In the second stage I will analyze the interactions between the texts to distinguish common features concerning my aim and research questions, and construct themes deriving from those features. The themes will lay the foundation for the presentation of my results. Throughout this work, I will analyze what the findings manifests in terms of power relations.

The last stage consists of contextualization; analyzing how the discourses revealed in the policies relate to socio-political context and discourse. Here I will also take social practices into account by also discussing the interface and device sensitive placements of the policies online as shown in web browsers and applications. Stanfill (2014) describes how a discursive analysis of interface aspects can be made. What I will deploy from this is the analytical aspects of the naming or labeling by which the user has to navigate to reach the policies, and the page placement in terms of placement of the links, buttons or menus which leads the user to the privacy policies (Stanfill, 2014, pp. 5-6). I will do this by analyzing these aspects as shown on a computer in the web browser Safari (before and after logged in to the service, except for Snapchat since it is not possible to log into the service through web browsers) and in the service's mobile phone applications for iOS/iPhone (except for Ello since the service is not available in the form of a mobile phone application). The goal with this analysis is to understand the online interface structure of the policies and the navigations to the policies as parts of the social determinations of the process of interpretation, and one aspect of the power structures surrounding the policies and how the power relations relate to the social-political context of the "big data society".

4.5 Summary and Considerations

By choosing to analyze the companies' policies and how they can be found I also limit the study to their words and their online context. The policies are written by the companies and are intended to be published for the public to read, as well as they are placed in an interface structure by the companies. It is reasonable to assume that the companies want to present a positive image of themselves in the policies, while reaching the demand (or at least give the impression of reaching the demand) for transparency. I believe it is important to bear in mind that the policies contain the perspectives and viewpoints of the companies, with a commercial agenda and an image to care for. At the same time, part of what calls for a scrutiny of the policies are these basic preconditions and how they shape the discourses of privacy policies and by them the notion of privacy in itself.

5. Description

By analyzing the privacy and data policies and their context in the online interfaces in accordance with my analytical approach, I distinguish and construct different themes. These themes and their titles derive from common features that are found in the empirical material, in accordance with my aim and research questions. The themes constitute different discourses that are interwoven and co-occurring, within the discursive field of the privacy and data policies. In this chapter I describe the results of the study, divided into the different discourses and themes I have distinguished. First I present the results of the interface analysis and then I continue with the results of the analysis of the privacy and data policies. The themes are called: Bottom Placement for Privacy, The Responsible User, The Good Deed of Collection, Sharing is Caring, "We may", The (Illusion of the) User in Control, and "Your Privacy is Important".

5.1 Interfaces

5.1.1 Bottom Placement for Privacy

When visiting the web sites of the social media services the navigation to the privacy policies is in almost all cases clearly expressed in terms of naming and labeling. The link stating "Privacy" (Facebook, 2015a; Google, 2015a; Instagram, 2015a; Snapchat, 2015a; Tumblr, 2015a; Twitter, 2015a; Youtube, 2015a), "Privacy Policy" (Linkedin, 2015a) or "Terms and Privacy" (Pinterest, 2015a). These labels clearly point to the content that the link refers to and is reasonable in terms of what the user will look for if she/he has concerns regarding her privacy. The naming of the link, "Privacy", does not change when logged in to the service in the web browser, though the placement in some cases changes. In the iOS applications of the services the user often has to go through menus, but one can argue that the naming still is clear, going through labels as "Settings" and then find "Privacy".

There are two apparent exceptions to this clearness of names though, and that is the labeling the user has to go through on Reddit's and Ello's web sites. Since the link stating "Privacy Policy" (at the absolute bottom of the long start page) is included in a sentence, very much toned down in color and in so small fonts that I first could not see it, I navigated through the menu to find the policy (Reddit, 2015a). The privacy policy can then be found through the link named "Wiki" and once on that page one has to scroll to the bottom and there "privacy policy" is under the headline "Boring Stuff" (Reddit, 2015b). When logged in to the service on Ello's web site, the user has to go through a menu called "WTF", and there find "Policies" which directs to the

privacy policy (Ello, 2015b). Whether or not these two examples represents accurate labels for the privacy policy to be under is arguable, but either way the labels and names of the links one has to go through to find the policies are rather unclear and not reasonable in terms of what labels the user will look for when having privacy concerns.

Compared to the mostly distinct and clear naming, the situation is reversed when it comes to navigation and placement. In the absolute majority of cases the link is placed at the absolute bottom of the page and in small, light fonts when the user is not logged in to the service (Google, 2015a; Pinterest, 2015a; Snapchat, 2015a; Tumblr, 2015a; Twitter, 2015a). In some cases it is also placed below the fold, which means that the user has to scroll down for the "Privacy"-link to be visible (Facebook, 2015a; Instagram, 2015a; Linkedin, 2015a), sometimes this means that the user has to scroll down really long, more than three times the length of the start page (Reddit, 2015a; Youtube, 2015a). Ello is one exception, with the placement of the "Privacy"-link at the upper right, next to, and as big as, the "Login"-link (Ello, 2015a).

As previously mentioned, in most cases the placement somewhat changes after the user has logged into the services. In some cases one has to scroll down a bit and the "Privacy"-link can be found in a menu at the lower right corner, in small light fonts barely standing out from the background, and never at the top of that menu's list of headlines (Facebook, 2015a; Linkedin, 2015a; Tumblr, 2015a; Twitter, 2015a). In other cases the user has to scroll down a great deal when logged in, to reach the "Privacy"-link at the bottom (Instagram, 2015a). The menu the user has to navigate while logged into Ello has already been explained, the placement although also differs. The "WTF"-menu is placed down in the left corner, in small fonts (Ello, 2015a). While logged in to Pinterest the user has to navigate a drop down-menu at the upper right corner and find "Privacy" at the bottom of its list (Pinterest, 2015a). Google, Youtube and Reddit proposes no changes in placement when logged in to the service (Google, 2015a; Reddit, 2015a; Youtube, 2015a).

When logged in to the iOS applications of the services the policies are generally harder to reach, as more steps are required of the user. In most cases the user has to go to their profile, find settings and find the policy there after scrolling down/made her/his way down the headlines on the settings menu, since "Privacy" never is among the top headlines (Google Inc, 2015b, 2015c; Instagram Inc, 2015b; LinkedIn Corporation, 2015b; Pinterest Inc, 2015b; Snapchat Inc, 2015b; Tumblr, 2015b). At its highest it is the third headline from the top (Tumblr, 2015b), and at it lowest "Privacy" is at fifteenth place (Instagram Inc, 2015b). To reach the policies in the applications of Twitter and Reddit it is a bit harder still. In Twitter's and Reddit's applications one has to go to settings, scroll down to the bottom of the headlines, press "About", and in the "About"-menu find the link to the privacy policy at the bottom (Reddit, 2015c; Twitter Inc, 2015b). Facebook's application has a slightly different structure and the user can, after she has pressed "More" in the main menu scroll down one extensive menu (how many headlines down depends on how many applications/groups etc. the user has) and at the bottom press "Terms & Policies"" and at that menu find a link to "Data policy". One can also reach it through "Privacy Shortcuts" in the "More"-menu, a headline that is slightly higher up than "Terms & Policies". In the bottom of the menu of "Privacy Shortcuts" the user finds the

headline “Data Policy”, although its appearance differs from the headlines above and it does not look like a link (Facebook Inc, 2015b). In the applications the user also has to be accustom to the symbols, which indicates the user’s personal profile (often some kind of figure resembling the head and shoulders of a human) or settings (often some kind of figure resembling a cog wheel), to reach the menus linking to the policies.

To conclude, the interface analysis shows that the naming and labeling of links to the privacy and data policies in most cases are clear and obvious, simply stating “Privacy”. However, the placements in the structure of the interfaces do not demonstrate the same clearness. The links are often placed at the bottom of the start pages in the web browsers and often below the fold. When logged in to the services the user often has to scroll down even further. The links are often of design features that do not stand out. In the mobile phone applications, the link to the privacy policies are mainly placed in the bottom of settings menus, and require more steps of the user to be found.

5.2 Policies

5.2.1 The Responsible User

The policies mostly refer to the user as “you” (your) and the service/company as “we” (us/our). In the policies, it is frequently formulated as if the information about the user is something that the user, deliberately and fully aware, chooses to give to the service/company. This is expressed by using wordings such as “you provide us”, “you send us” or “you give to us”. This is very much apparent in for example a headlines used by Snapchat: “What You Directly Provide Us” (Snapchat, 2015c), and a description used by Pinterest:

> When you sign up for or use our products, you voluntarily give us certain information. This can include your name, profile photo, Pins, comments, likes, email address you used to sign up, and any other information you provide us.
>
> *Pinterest, 2015c.*

This makes the user the agent in those sentences, which implies that the user is the active part that is responsible for the events/actions explained in the policies. By not writing “we take/gather this information from you”, and instead stating “you provide this information”, a type of passivization regarding the companies deeds is made. There are numerous examples of this in the privacy policies, such as “Information you provide us directly: […] User content (e.g., photos, comments, and other materials) that you post” (Instagram, 2015c). In some cases even the “us” (referring to the service/company) is absent from the sentences, as in: “When you create or reconfigure a Twitter account, you provide some personal information, such as your name, username, password and email address” (Twitter, 2015c). By the use of the verbs “provide” and “give” one could also argue that these works in euphemistic ways, to make the actions seem more positive.

In the previous examples the information "provided" is rather straightforward, often by such as formularies that the user has to fill out, but there is frequently use of this formulation also concerning information that is not so deliberately provided by the user. It is portrayed as if the user choose to give the information to the service, in apparent ways like stating one's name and e-mail address but it also works in less apparent ways. By using the service the user is viewed as deliberately giving the information that is gathered. For example, Snapchat has another headline stating: "What You Automatically Provide Us When You Use Our Services" (Snapchat, 2015c) and Google writes: "We collect information in two ways: Information you give us […]. Information we get from your use of our services." (Google). The formulations in the policies are emphasizing the awareness and responsibility of the user:

> By going to those links or by using a co-branded or third-party-branded Service, you may be providing personal information directly to the third party, us, or both. You acknowledge and agree that we are not responsible for how those third parties collect or use your information.
>
> *Snapchat, 2015c.*

One headline in Instagram's policy states: "Parties with whom you may choose to share your User Content". This refers to photos that users choose to publish, but the policy also highlights that this choice also entails what publishing the photos on the Instagram API means: it becomes available for third-party companies to use. By choosing to use the service the user, according to the policies, willingly discloses the information and make the service/company co-owner of that information. It implicates consent to give up ownership and puts all responsibility onto the user.

> By providing personal information to us when you create or update your account and profile, you are expressly and voluntarily accepting the terms and conditions of our User Agreement and freely accepting and agreeing to our processing of your personal information in ways set out by this Privacy Policy.
>
> *Linkedin, 2015c.*

The service as passive part in the transaction of information is further emphasized in some of the policies, which frequently phrases that the service/company is "receivers" of information.

> You allow us to receive information when you use your account to log in to a third-party website or application. Also, when you visit a third-party site that embeds our social plugins […] we receive information that those pages have loaded in your web browser.
>
> *Linkedin, 2015c.*

This implies that the service/company is the passive part, not responsible for being "provided with" and storing the user's information. By taking the deed of the social media companies out of the equation it gives the impression of the act as non-existing, of the services as passive receivers of information. As Twitter writes: "You may also tell us your location when you set your trend location on Twitter.com or your computer or mobile device sends us location information" (Twitter, 2015c). Although it continues with describing other ways they may use other data from the user's device to determine the location, this first sentence signals that this information

is sent to Twitter whether Twitter like it or not. It removes the conscious action of Twitter to gather, store and use that information.

This is also apparent in one recurrent word in the policies: automatically. This also suggests that the action/event described is out of the service's/company's hands.

> We automatically receive and record information from your web browser when you interact with the Services, such as your browser type […] what sort of device your using, […] your language preference, the website or service that referred you to the Services, the date and time of each request […], your screen display information, ad information from any cookies we have placed on your web browser.
>
> *Tumblr, 2015c.*

> These days, whenever you use a website, mobile application, or other Internet service, there's a certain information that almost always gets created and recorded automatically. The same is true when you use our products.
>
> *Pinterest, 2015c.*

This quote from Pinterest's policy uses automatically in a way that signals that the company is not in any way blamable for creating and recording that information. This process is described as something that "just happens", almost as if it suggests that it is part of a natural ecosystem. This use of "automatically" is a part of a frequent description of technology as responsible actor, next to the user, in the policies, such as in: "your computer or mobile device sends us location information" (Twitter, 2015c).

Over all, by using the wording *you* provide us or *you* allow us, the focus is shifting from the service/company as an actor and onto the action made by the user. One part of this is that the companies actions, if they are described as something that the company does, is formulated as motivated by the user's actions and aware decisions, as in: "we may share information about you with business partners to provide the Services and functionality you request and to communicate with you about those Services" (Snapchat, 2015c). By writing *you request* the company place the action onto the user. Another example of this is from Facebook's policy:

> Depending on which Services you use, we collect different kinds of information from or about you.
>
> Things you do and information you provide.
>
> We collect the content and other information you provide when you use our Services […]. This can include information in or about the content you provide, such as the location of a photo or the date a file was created
>
> *Facebook, 2015c.*

It is also expressed as if the companies have no choice but to gather this information and to require the user to give up certain information, such as name and email address, to be able to create an account and fully use the service. One exception to this is Reddit, that do not demand the user to provide her name and to provide an

email address is optional: “When you create an account, you are required to provide a username and a password, and may opt to provide an email address” (Reddit, 2015d).

As a whole this use of language suggests that the user-generated data are, consciously, given from the user to the services. This can be seen as a way to make the user, and not the services/companies, solely responsible for this transaction.

5.2.2 The Good Deed of Collection

> We may use this information about how you and others interact with the Services for a number of things generally related to enhancing, improving, protecting, and developing new Services, including but not limited to: providing users with personalized content; providing users with targeted advertising
>
> *Tumblr, 2015c.*

The services’ gathering and constructing of information about the users and the users’ behavior is in the policies throughout called “collecting”. To “collect” can in itself be considered a choice of word to make the action seem more positive, since it constitutes more positive connotations than to gather, compile and construct information about the user, which implies more of a purpose for use of that information. Surrounding this “collection” that is done by the services/companies, are other euphemistic uses of words. One frequently recurring word is improve.

The services/companies motivate much of the collection of information by stating that it is used to improve the service and the experience of the user, such as in: “Twitter uses Log Data to provide, understand, and improve our Services“ (Twitter, 2015c) and “We may use Account Information, alone or together with other information, to enhance and improve the Services, such as by personalization” (Tumblr, 2015c). This collection is portrayed as something the company does as a favor to its users, as when Snapchat in its policy answers the question of what the service do with the information that is collected: “The short answer is: Provide you with an amazing set of Services that we relentlessly improve” (Snapchat, 2015c). The same is true regarding the personalization of information made possible by this gathering:

> Twitter may keep track of how you interact with links across our Services […] We do this to help improve our Services, to provide more relevant advertising, and to be able to share aggregate click statistics such as how many times a particular link was clicked on
>
> *Twitter, 2015c.*

The word ”improve” is not always used, but it is still emphasized that the collection is done to make the service better and as something that is positive for the user:

> We collect information to provide better services to all of our users – from figuring out basic stuff like which language you speak, to more complex things like which ads you’ll find most useful, the people who matters most to you online, or which YouTube videos you might like
>
> *Google, 2015d.*

This focus on improvement of the services can to some extent be seen as a truth with modification, since this "collection" of data is what makes the companies interesting for advertising agencies and thereby how the company earns money. In that sense collecting of different user-generate data do not solely aim to improve services. It is also worth noting that this shows that it is not just the collection which falls under this good deed for the user discourse, it is also the store and use of the information about the user, as previously mentioned regarding personalization and which is also apparent in for example Google's and Facebook's policies:

> Our automated systems analyze your content (including emails) to provide you personally relevant product features, such as customized search results, tailored advertising, and spam and malware detection
>
> *Google, 2015d.*

> We want our advertising to be as relevant and interesting as the other information you find on our Services. With this in mind, we use all of the information we have about you to show you relevant ads
>
> *Facebook, 2015c.*

The store and use of user-generate data is mediated as something positive for the user, for example, Instagram emphasizes that the service use information they "receive" to make it easier for the user by: "remember information so you will not have to re-enter it during your visit or next time you visit the Service" (Instagram, 2015c). LinkedIn also constitutes an example of formulating gathering, storing and using the user's information as a good deed to serve the user and make ads "relevant and useful" to the user:

> we use cookies and similar technologies, including mobile application identifiers, to help us recognize you across different Services, learn about your interests both on and off our Services, improve your experience, increase security, measure use and effectiveness or our Services, and serve advertising.
>
> *Linkedin, 2015c.*

The policies of Reddit and Ello do not emphasize collection, store and use of personal information as something mainly positive in the same way as the other policies do. However, their privacy policies are not free from these types of statements. Fore example, Reddit writes: "to avoid showing you the same ad over and over again, we share your device's unique advertising identifier" (Reddit, 2015d).

Through and out, the social media companies in their policies mainly depict the collection of user-generated data as something positive. This is done by stating that it is done to improve the service and make the user's experience better. Not showing the user the same advertisement and to provide the user with customize advertisement, is described as a favor of the companies.

5.2.3 Sharing is Caring

Another frequently recurring word in the policies is the word share. As the concepts previously mentioned, this also can be seen as a euphemistic use of the word share,

since to "share information" can be argued to entail with much more positive connotations than to "sell information" or use the user's information to make business deals/earn money. As formulated by Pinterest: "We may also share aggregated or non-personally identifiable information with our partners, advertisers, or others" (Pinterest, 2015c). To "share information" indicates that it is not something the service/company does to make profit, sharing rather indicates good will and hence can be considered used as a way to win acceptance for the actions of the company.

> We never share information we receive from you unless: (a) we have your permission to share that information: (b) we have given you prior notice that the information will be shared, and with whom (such as in this Privacy Policy); or (c) that information is aggregate information or other information that does not identify you.
>
> *Tumblr, 2015c.*

The majority of the policies do not firmly acknowledge that the user's information could be a commodity and that this "sharing" is a part of how the companies make their profit; it is mostly not mentioned as commercial goods except for when the companies states that if they were acquired by another company, users' information would be part of that business transaction. There are exceptions to this though. While they are not free from using the word share or sharing, Ello's and Reddit's policies address the concept of selling information. Ello emphasizes that the company doesn't make profit by using the user's information:

> Ello does not make money from selling advertising on the site, serving ads to you, or selling information about our users to third parties, including advertisers, data brokers, search engines, or anyone else.
>
> *Ello, 2015c.*

Similar statements are made by Reddit, accentuating that "Your Private Information Is Never for Sale" (Reddit, 2015d) and:

> While advertisers may target their ads to the topic of a given subreddit or based on your IP address, we do not sell or otherwise give access to any information collected about our users to any third party.
>
> *Reddit, 2015d.*

As in the previous quote from Tumblr's privacy policy, the most policies states that the information they "share" with third parties is non-private information, non-identifiable information, information which they claim can not be traced back to a specific individual: "We do not share your personal information with any third-party advertisers or ad networks for advertising without your separate permission" (Linkedin, 2015c). What information Linkedin classifies as "personal" though remains unclear to the user and in most policies it is hard or nearly impossible to figure out what is included in the scope of "personal" or "private" information.

> We may share or disclose non-private information, Aggregate information, or other non-personally identifying information with people and entities that we do business with
>
> *Tumblr, 2015c.*

Some of the policies include examples of what "personal" or "private" information could be:

> We may share or disclose your non-private, aggregated or otherwise non-personal information, such as your public user profile information, public Tweets, the people you follow or that follow you, or the number of users who clicked on a particular link […], or reports to advertisers about unique users who saw or clicked on their ads after we have removed any private personal information (such as your name or contact information)
>
> *Twitter, 2015c.*

> We do not share information that personally identifies you (personally identifiable information is information like name or email address that can by itself be used to contact you or identifies who you are) with advertising, measurement or analytics partners unless you give us permission.
>
> *Facebook, 2015c.*

On the basis of these and similar examples personal and private information consists of such as name and email address, while all other user-generated information not is considered personal or private. Facebook further gives an example of what they consider to be non-personally identifiable: "such as 25 year old female, in Madrid, who likes software engineering" (Facebook, 2015c). Google states that the company "may share aggregated, non-personally identifiable information publicly and with our partners – like publishers, advertisers or connected sites" (Google, 2015d). And also states that the company requires opt-in consent for "the sharing of any sensitive personal information" (Google, 2015d). What the company considers as non-personally identifiable or sensitive information is not explained in the policy. It though contributes links to a web page with key term explanations, where Google states that
non-personally identifiable information is information that "no longer reflects or references an individually identifiable user" and that sensitive information is a category of personal information "relating to confidential medical facts, racial or ethnic origins, political or religious beliefs or sexuality" (Google, 2015e). It demands an extra step for the user to go from the privacy policy to this web page and still, the explanations given by Google are very ambiguous and do not tell which specific information will be "shared". These kinds of uncertainties, which the users are left with, are to a large extent consistent in the policies.

To conclude, in the policies the word share is used to describe the process of giving access to user-generated information or in some way disclose it to third parties, even though this often constitutes business transactions. Exceptions are Reddit and Ello that state that they do not make profit out of user-generated data and they do not use the word share in the same way. In the policies it is often unclear what information will be "shared". When the companies write that they do not "share" private or non-identifiable information, what information they refer to is ambiguous.

5.2.4 "We May"

A phenomenon that constitutes a major part of the uncertainty-factor of the policies is the frequent use of "we may". This concerns the "collecting", storing, "sharing" and usage of the user's information, explained by the companies.

> We may remove parts of data that can identify you and share anonymized data with other parties. We may also combine your information with other information in a way that it is no longer associated with you and share that aggregated information
>
> *Instagram, 2015c.*

The word ”may” suggests that there is no way that the user can know if this action will happen and if so how. Does “we may” in that sentence, suggest that identifiable data maybe not will be removed, or that the combining of information not will be made, before “sharing”? The combining and removing do in themselves constitute uncertainties regarding what they purport, but the “may” poses the greatest insecurity in terms of what will and will not happen. The sentences often also consists of other uncertainty-element, such as “in some cases”, “certain information” or by mentioning examples that maybe just represents a small part of when or how an action can occur. Tumblr writes: “In some cases, we partner with Third Party Services that may provide information about you” (Tumblr, 2015c), in which the “in some cases” and “may” proposes double uncertainties. There are numerous examples of this use of language in the policies.

> We *may* let other companies use cookies, web beacons, and other technologies on Snapchat. These companies *may* collect information about how you use the Services and other websites and online services over time and across different services. The information collected *may* include unique device identifiers, […] links clicked, and conversation information. This information *may* be used to, among other things, analyze and track data, determine the popularity of certain content, and better understand your online activity [my italics].
>
> *Snapchat, 2015c.*

Recurrent formulations including “we may” makes it unclear to the users if, when and how information about them will be “collected”, stored, “shared”, used and removed, as well as what information these actions will concern. In the example from Snapchat’s policy above, every sentence has a “may” in its beginning, which illustrates how this recurrently is used in the policies. Reddit represents an exception with very few “we mays”. Much information in Reddit’s policy is stated clear, or more clearly, than in the other companies’ policies, such as “by default, [post and comments] are not deleted from our servers – ever – and will still be accessible after your account is deleted” and that Reddit “stores the IP addresses associated with specific posts, comments, and private messages for 90 days after they are made or sent” (Reddit, 2015d).

The frequent use of “we may” is a way of telling that the user should count on this to happen, but it is not sure that it will. Probably this is a way for the companies to guard themselves, to be able to say “may” so it seems like something that is not happening all the time, or if it is happening the company can claim it informed its users about it in the policy. But in reality it could also propose a power hold, leaving if and when certain information will be collected, used and stored unclear for the user.

5.2.5 The (Illusion of the) User in Control

Throughout the policies the information is described as *your* information, the user's information. But still the rights of ownership in terms of control over the information do not seem to remain in the hands of the user. In the privacy policy of Google, one can read:

> Whenever you use our services, we aim to provide you with access to your personal information. If that information is wrong, we strive to give you ways to update it quickly or to delete it - unless we have to keep it for legitimate business or legal purposes.
>
> *Google, 2015d.*

It is called *your* personal information but still the access to it is not undeniable and is something that Google *aim* to provide. That the user should be able to modify the information is something that Google *strive* to enable. These actions are presented as favors to the users. Is the information really *yours*?

From the moment that one becomes a user of the services onwards it is, in the policies, portrayed as if the user is the one that has control over her information and that every bit of information that is "shared" is done so by informed choice. As discussed above the personal information is portrayed as something that the user chooses to give away. In the beginning of Twitter's privacy policy it states that:

> When using any of our Services you consent to the collection, transfer, manipulation, storage, disclosure and other uses of your information as described in this Privacy Policy.
>
> *Twitter, 2015c.*

The focus is, as previously mentioned, the user and the user's action, and not what Twitter actively are doing or aiming to do. This quote once again illustrates this. The previous sentence is followed by another statement, which removes Twitter as agent and portrays the user as in control:

> This includes not only the messages you Tweet and the metadata provided with Tweets, such as when you Tweeted, but also the lists you create, the people you follow, the Tweets you mark as favorites or Retweet, and many other bits of information that *result from your use of the Services*. [my italics].
>
> *Twitter, 2015c.*

The policies also to some extent imply that the user can choose not to allow collection of certain data, but one could claim that this choice seldom exists in practice, if the user still wants to be able to use the service. This is common regarding the use of cookies and similar technologies, and the user's ability to not allow cookies. For example, Instagram states about the user's ability to not allow device identifiers: "Some features of the Service may not function properly if use or availability of device identifiers is impaired or disabled" (Instagram, 2015c). This suggests that it is just an illusion of choice, if the user should continue to use the service. Google states that the company's goal is to be clear regarding what information the company collects; "so that you can make meaningful choices about how it is used" (Google, 2015d). Google then lists examples of settings the user can modify, such as view and

edit preferences about what ads are shown to the user and opt out of some advertising, take the information that is associated to one's Google Account out of some of the company's services and choose if one's profile name and photo can appear in ads. But still this is just small pieces of how the information is used and the user cannot opt out of the "collection" itself. Google mentions that the user can block cookies in her/his web browser, but follows with: "However, it's important to remember that many of our services may not function properly is you cookies are disabled" (Google, 2015d). This type of reasoning is apparent to very large extent in the policies, and illustrates that what is presented as a choice seldom constitutes a real choice:

> With your consent, we collect your device location so that you can use our location-based features, […] you can always revoke your consent by changing the settings or preferences on your device. If you do so, certain features of the Services will no longer function.
>
> *Snapchat, 2015c.*

Reddit and Ello though propose some exceptions to this. For example Reddit states that they want to make it easy for the user to opt out of sending her device's unique advertising identifier to the company and provide a link to instructions on how to opt out, a process which is fairly easy (one simple and distinct opt out-button in the application's settings) (Reddit, 2015d). Ello make up an exception by allowing the user to opt out of analytic tools, although the formulation that Ello will "make best efforts" is rather ambiguous:

> On your Ello settings page, you can choose to turn Google Analytics off completely when you visit the Site. If you choose either of these options, we make best efforts not to send any data about your user behavior, anonymized or otherwise, to Google or any third party service provider
>
> *Ello, 2015c.*

Most of the times though, the user is given no choice at all. The events and actions are just stated as something that will, or "may", happen, and the user is considered to have agreed to those actions, due to her use of the service:

> When you use Google services, we may collect and process information about your actual location. We use various technologies to determine location, including IP address, GPS, and other sensors that may, for example, provide Google with information on nearby devices, Wi-Fi access points and cell towers
>
> *Google, 2015d.*

The privacy policy of Twitter (2015c) advises users to think carefully before tweeting since tweets immediately are sent by SMS and API to third parties. That means that after the user has posted content, it is not really possible for the user to ever fully remove that content. In other words, even if the user later can remove the tweet, it has already been stored elsewhere. The same is true for Instagram:

> If you remove information that you posted to the Service, copies may remain viewable in cached and archived pages of the Service, or if other Users or third parties using the Instagram API have copied or saved that information
>
> *Instagram, 2015c.*

While the policies claim that the user is in control over her information, *your* information, this illustrates how the user loses control over her content after posting it to the services.

The (insignificant) room for action the user has is further apparent in the companies previously mentioned ambiguously or non-existing definitions of what personal and non-personal, as well as non-identifiable, information is. This is something that the companies decide and that the user has no control over. When it is not clearly explained what that information is, it is also hard for the user to make informed choices of information to disclose or not. The policies also show that they do not only concern the users of those specific services, but everyone who visits websites that include plugins from the companies' services: "We collect information when you use your account to sign in to other sites or services, and when you view web pages that include our plugins and cookies" (Linkedin, 2015c). This further diminishes users' control over their information, since they may not even be aware of the fact that they are objects of the companies "collection" of information.

Consistently the information about the user is addresses as *your* information by the social media services' privacy policies, as to signal that information about you, is yours. More truthfully this information should be called *our* (the service's) information *about you*.

5.2.6 "Your Privacy is Important"

Throughout, the policies refer to privacy as something that the social media companies hold high, something that is of great importance. As Tumblr (2015c) states: "Tumblr, Inc. […] takes the private nature of your information very seriously." Similar formulations can be found in all of the policies. Linkedin (2015c) even uses the title "Your Privacy Matters", instead of simply use "Privacy Policy". Reddit (2015d) takes it a bit further and states that "your privacy is genuinely important to us.", which suggests that in comparison to others claiming the same, Reddit actually means it. Also Ello is keen to stress that the service's handling of privacy related aspects stands out:

> Ello […] takes data privacy seriously. As a network that does not serve advertisements and that does not sell information about its users or use of the Site […] to third parties, Ello has taken unique steps that help you control how much information about you is shared when you use the Site.
>
> *Ello, 2015c.*

Privacy is depicted as something that is valuable to the user, and as something that *belongs* to, and is in possession of, the user. The user could have privacy (*your* privacy) and control privacy. At the same time as this is depicted as highly valued and of great importance, the concept of privacy is discussed as something that can and is threatened, or feared to be threatened, by parts of the services. The reasoning proposes a paradox; the services/companies claim to consider the user's privacy as important, and at the same time represent a service that entails threats to the user's privacy. This constitutes the base for the existence of the policies and is often

addresses in some way or another at the beginning of the policies. For example Twitter (2015c) writes: "We do not disclose your private personal information except in the limited circumstances describes here." In some cases the companies claim they cannot take fully responsible for the user's privacy when using the service, in terms of for example what other users might do with the public information, or other more far-reaching statements. Instagram (2015c) claim that the company: "cannot ensure the security of any information you transmit to Instagram or guarantee that information on the Service may not be accessed, disclosed, altered, or destroyed." In the same manner, Reddit (2015d) states that: "no data transmission over the internet is completely secure, so we cannot guarantee the absolute security of this data. You use the service at your own risk." This can appear alarming but at the same time maybe constitutes a more truthfully statement than to claim otherwise.

At the same time as the companies claim to prioritize privacy and highlight how important they consider it to be, aspects of the policies suggest otherwise. One very clear example is Tumblr, which according to the previous quote claim to take the user's privacy very seriously - something that is contrasted by the statement:

> Tumblr may determine your location by using drone technology and live video feeds. Ha ha, no, we just check out your IP address or any location data you attach to a post. Normal stuff.
>
> *Tumblr, 2015c.*

The fact that Tumblr in its privacy policy jokes about how location data is gathered do not suggest that the company takes their user's privacy concerns very seriously. One could also wonder whether the joke about extreme methods is a way to gain acceptance for the methods they actually use, by making them seem less extreme. Some of the policies also excuse the use of technical language, claiming that they try to make it easy and simple, for the user to understand. As Pinterest (2015c) states: "Because we're an internet company, some of the concepts below are a little technical, but we've tried our best to explain things in a simple and clear way."

Many of the policies refer to other policies. Both other policies of the own company, such as cookie policies, and other companies' policies, such as analytical services of other companies that the service uses, which they consider that the user should read as well.

> Among the third parties we use for analytics is Flurry, which provides mobile analytics in connection with our app. For more information about Flurry's privacy practices, or to opt out of Flurry Analytics tracking through our app, please visit www.flurry.com/user-opt-out.html.
>
> *Snapchat, 2015c.*

This means that in many cases, reading the companies' privacy or data policies is not really enough to fully comprehend the gathering and use of one's information. It could also be hard to determine if features of the service belong to the company of the service or a third part with other privacy principles. This is especially apparent on Facebook, where the user can deploy an extensive amount of third-party applications:

> when you download and use such third-party services, they can access you Public Profile, which includes you username or user ID, your age range and country/language, your list of friends, as

> well as any information that you share with them. Information collected by these apps, websites or integrated services is subject to their own terms and policies.
>
> *Facebook, 2015c.*

To conclude, one could also argue whether privacy is depicted as a property, capacity, quality or state of being. And if the policies concern themselves with an actual state or rather a feeling of a state, as well as the user's actual control or rather the user's belief or feeling of being in control. Through and through, the policies start by describing privacy as something valuable and that the privacy of the user is important to the company. However, these claims do not correspond with later statements by the companies that do not depict privacy as a priority.

5.3 Summary

I see the themes elaborated in this chapter as representing different discourses, manifested by the privacy and data policies of the social media companies. The discourses are in somewhat overlapping and some of them are more distinct than others. The seven discourses are all part of the discursive field of privacy and data policies and in different ways imply power structures regarding ownership and control of user-generated data, as well as regarding the construction of privacy. In them there is also signs of discursive struggle, as dominant discourses are contested by opposing statements. I continue this analysis by discussing these discourses in relation to each other and in terms of my research questions and research aim.

6. Interpretation

The themes elaborated in the description represent different discourses and different aspects of the privacy and data policies and their online context within the interfaces. In various ways these discourses express perceptions of ownership and control of user-generated data and the user's room for action, as well as privacy. In this chapter the results are further analyzed in relation to previous research and the theoretical framework, in accordance with my research questions and my research aim.

6.1 Construction of Privacy

The mechanisms of labeling the policies as *privacy* policies, how the word privacy is used in the policies and their online context, is similar to the process of naming analyzed by Olson (2002). The policies are part of creating the "identity" of privacy, the *construction* of privacy: what it consists of, how it should be perceived and properly managed. The formulations of the policies are ways to achieve naturalization of discourses and to impose their views on privacy as commonsensical, as in Fairclough's definition of ideological power. As Fairclough (2001, pp.78-89) argues, the meanings of words are not fixed and are affected by discourse, in a process where the struggle over language can end in a discourse undergoing a fixing/closing of meaning, as it becomes naturalized/commonsensical. Fairclough (2001, p 38-39) argues that power *in* discourse is powerful participants controlling the content and contributions from non-powerful participants. As with traditional media he refers to the one-sidedness of it. The sharp divide between producer and interpreter is also true for the social media services in this study. The policies and their contexts in the interfaces are presented for the users, leaving them with the option to buy into it or to reject it, by not being part of it at all. The labeling of the policies as "privacy" policies signals that what is treated in it is what privacy consists of, and thereby the companies gets interpretative prerogative over what privacy is, what rightful threats against it are and how privacy is properly protected. The construction of privacy, and what rightful claims about privacy concerning user-generated data are, is what the policies and their online context are trying to establish. The results of my study shows that privacy is portrayed as valued, that the user should have some privacy concerns, but that the concerns are properly taken care of by the companies and that the user can control her own privacy as described by the policies. What is not mentioned in the policies is not a privacy matter, or at least not something important enough to be mentioned.

People's views of what privacy is includes secrets, protecting personal information and having control over one's information (Vasalou et al., 2014, pp. 9-11). Also Petronio (2002) includes sense of control as well as ownership of the personal

information in the definition of privacy by CPM. The results of this study suggest that, in the constructions of privacy in the policies, the *ownership* and *control* of the personal information is undermined. The companies begin with claims about the importance of the user's privacy, that the information belongs to the user and that she can control her information, but this stands in contrast to the statements and terms, which restrict ownership and control, that follows in the policies. At the core is what is called personal information and the control over it. In a similar way as the paradox of openness and closedness in Haider's (2008) study, the discourses of the policies proposes a Janus-faced notion of privacy. At the same time as the policies articulate that the companies care about and protect privacy, they originate from the state of the services as something that is privacy threatening. Technological determinism is also something that is present in the policies, as in Haider's (2008) study. Within the discourses of open access she finds a view of technology as untroubled "neutral facilitator of development" (ibid, p. 118). In the privacy and data policies this can be seen in the formulations of the companies as passive receivers, that technological development makes certain tracking, collecting and use possible, and hence the companies do it. But this can also be seen in the statements of actions happening "automatically", mediated as if the companies have no choice or no optional course of action. Technology is depicted as a responsible actor as the same time as the companies' agency often is left out the sentences.

The accentuated claim of importance of privacy, made by the companies in the privacy policies, is not mirrored in the placement of the policies in the interfaces of websites and mobile applications, as shown in the results. The privacy policies are not placed high up on the pages or in the menus and the user is often forced to scroll down to see the links, which often implicates that that information is overlooked (Stanfill, 2014, p. 6). The results do not agree with the "path of least resistance". My findings stand in contrast to the claims made in the policies, of privacy as something highly valued and emphasized, as it do not agree with how something valuable is presented in interfaces, where higher placements indicates visibility and the weight that is ascribed to that information (ibid, p. 6). Making something stand out and notable relates to how visible the service wants that information to be, and implies the valuation of that information. For example Debatin (2011) and Nissenbaum and Barocas (2014) states that few people read and are aware of privacy policies, and this may be an aspect of placement and design choices which do not encourage the policies to be read. The results agree with Jensen and Pott's (2003) findings, that the user has to actively be looking for the privacy policies to find them as one are not presented to them otherwise and are unlikely to follow links to policies. And the user presumptively needs to be concerned about these issues to seek out and find the privacy policies (Jensen & Potts 2003).

One of the aspects Fairclough (2001, p. 52) acknowledges as power *behind* discourse is who has access to the discourse and impact over giving access. The user has no access to powerful subject positions within the discourses, other than choosing not to use the service, which has no effect when just one individual does it. Neither has the user at her subject position the possibility to access the discourses in terms of ability to change the statements in the policies; they are non-negotiable (Jensen & Potts, 2003). The struggle to try to naturalize the discourses of privacy and user-generated data and posses the power behind discourse, is present in how the policies

corresponds with each other, or at least how a pair of them correspond with the rest. The policies of Ello and Reddit do this by stating that they *actually* care about their users privacy, and similar statements, made to declare that they are more genuine in their protection of the users' privacy than the other companies are, even though those policies also claim to care about users' privacy. In some aspects the policies of Ello and Reddit constitute a discursive struggle, where they want to be profiled as fighting the discourses of an illusive, non-truthful, claim to care about and protect the users' privacy.

6.2 Personal Information as Possession

When the policies of the social media companies state that it is *"your* information" and that "you provide us" or "you give us" that information, the "giving" establishes that the personal information is something that belongs to the user and which the user chooses to give away. This corresponds with the co-ownership concept of CPM. Petronio (2002, p.1) declares that we feel that "we are the rightful owners of information about us." This is a phenomenon acknowledged by the discourses of the policies, first of all by constantly calling it *your* information when referring to information about the user. The management of boundaries within CPM theory marks ownership lines and we are believed to be weighting risks against benefits before disclosing information and engaging someone in co-ownership (ibid, pp. 6-10). When someone else attempts to control information we *perceive* as ours, Petronio (2002, pp. 6-10) emphasizes that it is considered a violation of privacy. By the results of this study, it can be argued that the information about us ceases to be ours as soon as it comes in the hands of the social media services. The policies are eager to mediate an image of what CPM calls *benevolent* ownership (ibid, pp. 130-131); that the companies take the discloser's wishes into account before "sharing" the information with third parties. However, based on the way the policies are formulated: the uncertainties, the lack of real choice for the user after the personal information has been "collected" or disclosed and the companies' lack of transparency, they rather express a *manipulative* ownership (ibid, pp. 130-132). With all the restrictions of the user's ownership and control of her information, the companies dominate how the information is managed and solely decides over access to the user's information, and hence hold a power position where they decide whether the user's personal information should be "shared" or revealed. If the user is dissatisfied with how her information is handled, they have to persuade the companies to change their practice, which may be hard or impossible for an individual to achieve. The policies make clear that the companies can sell, transfer, rent and share user-generated data, whether or not this is considered morally correct it illustrates that the data about the users are no longer in the hands of the users. If the companies do not sell it, or do so with alterations they claim is de-identifying, it is portrayed as something that is done out of courtesy of the companies. Through and through, this shows that user-generated data is commodified. Fairclough writes that what a commodity is has "expanded from being a tangible 'good' to include all sorts of intangibles: educational courses, holidays, health insurance, and funerals are now bought and sold on the open market in 'packages', rather like soap powders." (Fairclough, 2013, p. 29). This is true also for personal information about individuals and this can be seen in the policies. Petronio (2002, pp. 191-192) states that the current "information explosion" makes

fuzzy boundaries common and that private information has become a commodity on the free market.

Within the discourses of the privacy and data policies, information is described as possession and hence agrees with Buckland's (1991, pp. 3-4) definition of information-as-thing. As long as the information stays as information-as-knowledge it belongs to the user and is in control of the user, but as soon as it is somehow expressed or turned into an action; by providing an e-mail address, publishing a Facebook-status, clicking on a link or visiting a website, it becomes information-as-thing and cease to be in the ownership and control of the user.

The one who discloses information considers herself as the original owner and by that, she thinks that she should be the one to determine any third-party disclosures of the information she co-owns with others (Petronio, 2002, p. 77). In the policies this is not the case, and it is often unclear exactly who these third parties that the information is "shared" with are and under what circumstances they will get the information. The frequent use of the word "share" is discussed by Williams, Agarwal and Wigand (2014), who argue that companies should stop hiding behind this word when they really mean "sell". Also Buchanan (2011) discusses the use of the term share as a way to disguise the profit-making transaction. I also find the use of the word as euphemistic, made to send a more positive message as an altruistic practice of an open connected web, and not acknowledge that it is a business transaction. The word "collection" is, as well as to "collect", used in a similar way. The companies are eager to by the policies establish this collection of data as a good thing - something that results in improvements for the user.

With all this in mind, the phrasing *your* information represents nothing but empty words in terms of who the owner is, who has the property of information in possession and can control and access it. The only way the information belongs to the user within the discourses of the policies is in terms of whom the information is about.

6.3 The User's Room for Action

Petronio (2002) states that central aspects of privacy are that people feel they own and want to control their private information. The policies express discourses that play on these aspects, eager to portray the personal information as belonging to the user and as if it is remaining in the user's control. The privacy boundaries pattern described by the policies is what CPM theory defines as *inclusive boundary coordination*, where the users give up control over their information to the social media companies (ibid, pp. 127-131).

The three boundary linkages types of this pattern, laid out in the theoretical framework, can all be found in the policies in different forms. Users are often forced to reveal private information if they want to become a user of the service, but at the other hand no one is forced to use the services, therefore the *coercive linkage* can be questioned. Although, people can have objections about the threats to privacy that a service could entail and still feel forced to use that service to be part of social life (Bechmann, 2014; Nina & Boers, 2013). We also have a need to be social as well as a

need to be private (Petronio, 2002, p.12). *Role linkage* (ibid, pp. 127-131) is highly present in the polices and their contexts as the companies hold power positions, dictating access to information after the user has disclosed information or, by the languages of the policies, agreed to collection of one's user-generated data. The user can have a hard time even getting access to data and information about herself, as this access is formulated as a favor to the user. *Susceptibility linkage*, when lack of self-monitoring puts an individual in a situation where she has disclosed more than her recipients, is apparent in the context of the policies but is not directly expressed though it is the premise for the platform – the users intentionally and unintentionally contribute personal information while the company often lack transparency. Since it is, as Nissenbaum and Barocas (2014) emphasize, nearly impossible to be informed about the scope of information that is gathered, one could also argue that self-monitoring online is nearly impossible.

The mediated hidden power described by Fairclough (2001, p. 43) can be found in how the policies are formulated. One aspect of this that is greatly present is how the policies express *causality*; whom that is represented as causing what is happening. When the companies writes "you give to us", "you provide us" or "you request"/ "you ask us", the companies express that the user is solely responsible for the action. This removes the agency of the companies and portrays the user as the agent. This is also true when the companies refer to themselves as mere receivers, which are not to be blamed for what they receive and how they receive it. As the policies are frequently using the phrasing "we may" there is a great uncertainty, which could constitute a base for *boundary turbulence* due to misunderstanding of the rules or fuzzy boundaries (Petronio 2002, p. 177). Since the co-owner, the user and the company, maybe believe in different rules and since few read the privacy policies, (Debatin, 2011), there is even greater risk for turbulence to occur. When the companies in their policies claim to highly value the users' privacy, the user could perceive this as a signal that her personal information will not be misused. As Jensen and Potts (2003) discuss, the fact that a website has a privacy policy is often misperceived as a promise not to sell information. Errors in judgment of what to disclose may also occur when people do not pay attention to rule development (Petronio 2002, p. 185). When the services updates their policies and the user also has all other services' and websites' policies to read, it is no easy task for the user to keep up with privacy rule development online.

The companies want to express that they do value privacy and that the users can control their own personal information. The policies claim that the user has room for action to control her privacy on the service. What the user actually can control is illustrated by the results of Stutzman and Kramer-Duffield (2010) of expectancy violations and privacy management, where the user can review what they post and be aware of the public information. The users seem less likely to disclose information after reading the policies, which suggests that the companies do not succeed with their image-making as privacy protecting. It could also be due to caution caused by the hardship of understanding the policies (Stutzman et al., 2010). People try to mend their disclosure because of privacy concerns, by less public disclosing and by changing their settings (Stutzman et al., 2012). But as long as companies like Facebook makes more information available to *silent listeners* (ibid); third parties, it is arguable whether the users has any control or if it is even possible for users to

advert the trend of increasing amounts of information to silent listeners. The policies states that preventative measures such as blocking cookies can be made, but that the service will not continue to be fully functional. It is as well hard to grasp the scope of the silent listeners. The policies do not make this easier and this suggests discourses in which real transparency is not promoted. This is somewhat contested by the policies of Reddit and Ello, for example by consisting of less uncertainties and the possibility to opt out of Google analytics tracking.

After consulting the privacy policy, the user can choose to not continue to use the service, or visit any websites with the service's plugins on, or opt out of Google analytics on Ello, but by then it is already to late since information already has been gathered before the user even has been be able to reach the policy. The user can be considered to be unaware of what third parties information will be "shared" with, even after reading the policies. As these policies relates to other policies as well, it is hard to distinguish which co-ownerships the user engage in. Also, with the "catch-22" situation in mind, explained by Jensen and Potts (2003), the user engages in the co-ownership and disclosure of information just by visiting the services websites or download the mobile applications, even before she has been able to read the policies. The companies in their policies puts the responsibility for this onto the user, but according to the discourses and social practices it is impossible for the user to control this, since no "safe areas" by which the user can reach the policies beforehand exist.

The user's control in the discourses of the policies constitutes merely an illusion of control, or a cosmetic control of public information. It is only controlling the user-generated data on the horizontal axis, the tip of the iceberg, while the data-collection on the vertical axis is to large extent invisible to the user and can not be controlled by the user of the services, since such as blocking cookies make the service not work fully and do not track-signals are not approved of (Debatin, 2011). How the policies stress the user's choice and control, but do not offer much of real choices if one wants to continue to use the service, the policies also makes it clear that what is described in them is, as Jensen and Potts (2003) states; non-negotiable.

The policies often make promises of anonymization and de-identification of the user's information, but it is unclear exactly what the policies refer to as they state this. What the companies call anonymization is most likely not enough to actually make the information impossible to trace back to individuals (Nissenbaum & Barocas, 2014). The unclear use of concepts like these is frequent in the policies. Fairclough (2001, p. 53) mentions literacy and digital literacy as aspects of access to discourse; can the users understand and comprehend what the policies describe? Jensen & Potts (2003) argues that many people do not have the required reading skills to do so, and even if they do, it is hard to comprehend what the privacy policies really entail and the scope of the actions described. As stated in some of the policies, they sometimes use technological terms that could be hard to understand. Nina and Boers (2013) also argues that few possess the media literacy skills to manage their privacy online. These statements can be considered as part of the discourse of "the information poor", as described by Limberg et al (2012). However, the discourses revealed in the privacy and data policies, in combination with previous research on readability and privacy policy consumption, calls for a greater discussion about privacy issues within media and information literacy education. However, it is not only a question of

readability; it also includes understanding of the online ecosystem, grasping the consequences of certain technological use online, and understanding of symbols such as the cogwheel indicating settings. In this sense, a broad set of media and information literacies are needed to make informed choices online. Olson (2002) discuss naming as something that could exclude and marginalize. By the discourses of the privacy policies, I consider that exclusion could occur at several different stages: in the digital/media/information literacy (finding policies), in the failing readability (understanding policies), shown by Jensen and Potts (2003), and also if a user do not agree and chooses to stop using the service (reject and protest policies). The question is whether there is somewhere for those privacy concerned people to go on the social web or if they become excluded from all online social communities, since they are not willing to give up their privacy. By increased understandings of these issues, whether or not it should be called privacy literacy or included in media and information literacy, users can become empowered. Literacy can be a key to transform society, as Limberg et al acknowledges (2012, p. 98), hence a key to transform how user-generated data is handled and improve social media users' ability to retain and control their own privacy online.

6.4 Power Manifested

Power by consent is increasing, as discussed by Fairclough (2001, p. 30), and he argues that people get integrated in "apparatuses of control" which people starts to feel a part of, as customers or otherwise. By the discourses expressed in the policies, describing the companies activities, social media services can be seen as such apparatuses of control.

Debatin (2011) stresses that privacy protection is becoming more vital as the capacities to process and store information expands when we move into the big data society. He further argues that people are unaware of privacy policies and settles for illusive privacy control rather than any real control. As mentioned earlier, Haider (2008) finds technological determinism in the discourses of open access, and such technological determinism is present also in the discourses of the privacy policies. The placements of the links to the privacy policies are very much alike from service to service, in a conformist way. This indicates technological norms of interface structures. However, this indication of technological determinism is most apparent in the numerous statements in the policies, which talk of things happening "automatically" and "nowadays". This indicates that the social media companies should have no other choice - because you can verify locations with Wi-Fi-connections you should also do so. This technological determinism is used to reduce the social media companies' role and responsibility over the actions and the technological development, it is depicted as something that just happens and is heading in a direction that the companies do not control. That it is bound to happen.

The policies states that the user consents to the actions laid out in the policies, but the premises are non-negotiable, which Jensen and Potts also discuss (2003). Since there is no real possibility for the user to change the premises and the companies strive to legitimize their power position, it becomes what Fairclough calls power by consent. Whether or not people likes what is described in the policies they join and use the services, which constitutes an image of that the collection, storing and using of

people's data is considered okay. How the policies of Reddit and Ello are formulated present, as previously discussed, a discursive struggle over the notion of privacy and the truthfulness of the statements of the value of privacy that is made. Fairclough (2001, p. 61) argue that the expressed aspects of discourse (power *in* discourse) constitute the site of the struggle, but what is really at stake is the ability to control the discursive premises (power *behind* discourse). This is done by establishing discursive domination. Although Ello and Reddit somewhat challenge the corporate discourses of privacy and power over user-generated data, they are not revolutionary different and they do not yet pose a threat to the other services and their actions. Moreover, their policies also include uncertainties and lack of transparency.

Privacy policy consumption decreases disclosure as people do not fully buy into the companies' claims about prioritizing privacy (Stutzman et al., 2010). At the same time, few read the policies and many believe the existence of a privacy policy entails guarantees for not selling the personal information (Jensen & Potts, 2003; Williams et al., 2014). Buchanan (2011) states that user's possibility to manage control is further reduced by external companies' entrance in the services, such as external companies' applications on Facebook, which users have to allow collecting information if they want to use them. External applications and services go by other privacy policies, as is stated in the privacy policies of the social media companies in this study. As Bechmann (2014) points out, the user then has to read all those privacy policies as well, and may not be aware that they go by other policies. I considered this as a power structure issue, and it is questionable whether these demands on users really are reasonable.

The results of the analysis of the policies and their interface context expose uncertainties and unclearness as the major power mechanisms of the policies, besides from obvious access and control. One aspect to consider when analyzing discourse is, according to Fairclough (2001) relational modality, and one distinct example of that in the social media companies' policies is the use of "we may". This express vagueness that can be used to the companies advantage and as a power hold; I see the vagueness as a way to exercise power, like Petronio (2002) discusses, fuzzy boundaries and unclear ownership causes turbulence. The fuzzy boundaries posed by the policies also stand in contrast to the idea of informed consent. In accordance with CPM theory, the user wants control over to whom and when the information is further disclosed. Therefore, the "original" owner; the user, should decide this and not someone who has been made co-owner of that information; the social media companies. The companies can argue that they will fulfill this wish, stating that they "may" "share" this and that "if" and "when" this and that happens, but the question is whether users can make decisions, and risk-benefit judgments, based on the policies, as they are filled with uncertainties such as "we may" and poorly defined concepts such as "personal information".

Petronio (2002, pp. 6-10) states that we agree to a contract of responsibilities when engaging in a co-ownership, but the vagueness of the policies constitutes a dilemma regarding this. Most do not read the policies, but for the one's that do it is hard to comprehend the implication of their content (Jensen & Potts, 2003). This relates to the idea of consent. As previous research shows (Bechmann, 2014), most people can not be considered informed and because of that, informed consent can be considered

a flawed method (Nissenbaum & Barocas, 2014). The premise for decisions and information practices on social media is the informed consent, but informed consent is a faulty principle and in reality may be impossible to accomplish (Bechmann, 2014; Nissenbaum & Barocas, 2014). The policies lack transparency in terms of elaborated descriptions of *when* information is gathered, *what* information it is, as well as *when* and *what* information is "shared", *how* and to *whom*. But as Bechmann (2014), Nissenbaum and Barocas (2014) propose, an elaboration of this, in terms of full transparency, would propose users with an unbearable load of information which most likely would be even harder to comprehend and users would be even more unlikely to read. This is truly a paradox and a dilemma, which calls for other privacy principles.

> There is little value in a protocol for informed consent that does not meaningfully model choice and, in turn, autonomy. The ideal offers data or human subjects true freedom of choice based on a sound and sufficient understanding of what that choice entails.
>
> *Nissenbaum & Barocas, 2014, p. 58.*

I suggest that by the discourses of the policies, in content as well as their interface aspects, they do not constitute a basis for informed choice. And they rather function, with the reasoning of Fairclough, as ways to legitimize asymmetrical power structures.

7. Explanation

The aim of this thesis is to investigate the power structures of privacy, ownership, store and use, of user-generated data, through the discourses manifested by data and privacy policies of social media services. We need bases for theorizing on how social inequity and power relations can influence information practices in LIS (Olsson 2010, pp. 67-68), and this is what I am hoping this thesis could provide to.

This study analyzes the social media companies' discourses of privacy and user-generated data, in the policies and in their place in the interfaces. It does not examine the user's reaction to these discourses, although the results are discussed in relation to previous research on users' privacy behavior, views on privacy policies and related aspects. The study concerns eleven services, and they are all in the category we call social media, I would however suggest that many of the patterns explored are general for the privacy policy context. Additionally, social media is more user-generated-data-intense than other services. Nevertheless, phenomena such as the transparency paradox are not specific for social media, as well as the power relation between provider (citizen), platform (companies and organizations) and analysts (advertisers and researchers) are not either. The discourse analysis offers a theoretical framework and method to reveal power structures and provided perspectives to this study that has not been discussed in previous research of privacy issues. By complementing the theoretical perspectives of discourse and power with CPM theory I contribute with a deeper understanding of mechanisms of privacy and at the same time provide further perspectives on how power and social structures can influence privacy issues. To complement the discourse analysis with interface analysis offered me an extended context and helped to pinpoint the interconnectedness of online structure as aspects of the discursive field of the privacy and data policies.

The fact that users react negatively on how the social media services handle their user-generated data and privacy, such as illustrated by the class action lawsuit against Facebook (Gibbs, 2015), demonstrates frictions between discourses. The lacking ownership and control aspects of the notion of privacy that the social media companies practice are not satisfactory to the users. The results of this study shows that the companies are eager to be seen as in favor of privacy protection and users' ownership and control of personal data, but a closer analysis of the statements they make paints us another picture. Nina and Boers (2013) argue that it is possible to change policies, by referring to how users were able to impact changes to Linkedin's privacy policy. However, I would argue that most users' discontent pass unnoticed. To initiate this sort of actions, lawsuits or major appeals, is also a David versus Goliath kind of game. By the time I write this it is still unclear whether the class action lawsuit against Facebook even will be taken to court.

Today, the scope of the social web stretches far beyond leisure use of social media. It is additionally a part of library service, supply and marketing. Privacy issues are also part of media literacy and information literacy. There has been no real change in privacy awareness during the last decade and education on privacy literacy is requested (Debatin, 2011; Williams et al., 2014). It is important to raise questions regarding online privacy, how privacy is constructed and what power structures that it is influenced by. As Burkell and Carey (2011) shows, libraries have much to improve to provide genuine privacy notice. When more of library service goes online and participatory web-technologies are implemented, will libraries fall in the same pitfalls as the social media companies, or will their democratic mission set other standards? The democratic mission and privacy protective code of ethics calls for the libraries to be good examples of this. According to Zimmer (2013), privacy issues are greatly overlooked in the implementation of new technologies within the library sphere. I believe if we want to continue to consider the library as a safe space, online privacy issues needs to be addressed to greater extent. And as library services move into the online participatory context social media's views on these issues needs to be taken into account.

Librarians are increasingly taking on the role of educators of citizens' on media and information related phenomena, which means that aspects such as privacy, identity, what is private and public, what information is gathered and how it could be used, needs to be discussed. Debatin (2011) does not relate privacy literacy to existing support for media and information literacy and do not acknowledge privacy literacy as part of those broader literacies. However, I view privacy literacy as a concept that should be included within information literacy and hence acknowledged and treated in media and information literacy education. Media and information literacy is vital to actively take part in, and be able to contribute to, society (Sundin & Rivano Eckerdal, 2014), and while a majority of the Swedes uses social media, awareness of its possible "hidden" consequences might not be as widely spread.

How Reddit and Ello express themselves maybe propose an alternative and another direction to the development of handling of user-generated data, especially so as Ello is a new service that profiles itself as valuing privacy in a more genuine way than other previous social media services. But skeptics do not believe that a business model that does not make profit out of personal information, and with no advertisements, will survive. Like discussed by Buchanan (2011), people are denied participation in social media if they are not ready to give up parts of their privacy; ownership and control over their personal data. I believe that it is important to keep in mind that the user-generated data of the social medias are used in many different contexts today. For advertising, market research and lobbying purposes but also for (public funded) research and similar areas.

The reason why Ello and Reddit represent these exceptions can maybe partly be connected to what their services are used for. Reddit is a discussion forum and do not need the social profiling and identification of a social networking service to function, the subject of the discussions and the discussions in themselves are what is important. By directed advertising subject-wise instead of pointing to specific individuals Reddit has not made itself dependent on user profiling for profit. Ello on the other hand wants to be an alternative to social networking services such as Facebook, with

personal profiles. As long as Ello does not have advertising or similar business models the user profiling is not a goal to reach profit, hence it as well is not depending on it for profit. There is a discursive struggle, but as long as services as Ello do not come near the amount of users that Facebook or the like have, the aspect of making profit out of the user-generated data is in dominance/hegemonic position in the discourses and services as Ello do not really pose as a threat or alternative to them as social platforms. While the big data-gathering expands, these issues can be expected to grow and the lawsuits or other raised voices could increase, but an individual have little to set against large companies with lawyers and policies shaped to disclaim the companies from responsibility.

The demands that the system of today puts on the user are impossible to live up to. To be informed of the policies of all social media services, their external partners and external applications, as well as the policies of all websites one visits and applications one downloads. To keep track of all privacy policies and changes made to them is an impossible task for the average Internet user. While efforts to support privacy literacy, and media and information literacy in general, are needed, these though are measures that point at individuals. Like the quote of Vaidhyanathan at the beginning of this thesis illustrates, the issues discussed in this thesis needs to be addressed as social issues – and not as individual issues. The emerging big data-society and the interconnectedness of today's online activities make it extremely hard, even impossible, for individual users to be informed of how all services, websites and companies handle user-generated data, according to Nissenbaum and Barocas:

> Typical of the big data age is the business of targeted advertising, with its complex ecology of back-end as networks and their many and diverse adjuncts. For individuals to make considered decisions about privacy in this environment, they need to be informed about the types of information being collected, with whom it is shared, under what constraints, and for what purpose. […] Simplified, plain-language notices cannot provide information that people need to make such decisions. The detail that would allow for this would overwhelm even savvy users because the practices themselves are volatile and indeterminate as new parties come on board and new practices, squeezing out more value from other sources of information (e.g. social graphs), are constantly augmenting existing flows.
>
> *Nissenbaum & Barocas, 2014, pp. 58-59.*

This quote illustrates that personal information and user-generated data are valuable recourses that are of high economical interest. This means that social media services and other user-generated data intense operations becomes powerful actors in the socio-political landscape, and even more so with the development of big data and its increasing possibilities to make value out of user-generated data: "Big data have many elements of a natural resource, and sensible rules must be developed in order to avoid a tragedy of the commons" (Lane et al. 2014, p. xiii). Accordingly this calls for action on a political, law-making level but actors with a mission to protect privacy and the interests of the public, such as libraries, also needs to speak up to make this happen.

Through and through, the ideological struggle that takes place in the language of the policies and their online contexts illustrates the power relation between company and user, and the larger power structures that controls the user-generated data. To control

the notion of privacy, and the companies' actions as enough to protect privacy, are ways to attempt to make these practices commonsensical and the companies free from responsibility. By the words of Fairclough: "The stake is more than 'mere words'; it is controlling the contours of the political world, it is legitimizing policy, and it is sustaining power relations" (Fairclough 2001, p. 75).

7. 1 Conclusions

The social media policies state that the users' privacy is something important and highly valued, however, this does not agree with the placements and design features of links to the privacy and data policies in the website and mobile application interfaces, where links to the policies mainly are placed in the bottom of pages and menus. In the policies privacy is constructed mainly as a possession, which the users choose to give to the services. By phrasing it as "you give" or "you provide us", the companies make the user the agent and themselves as passive receivers. This is a way to express causality that makes the user seem like the one solely responsible for what is happening. The policies are trying to depict an image of the user as the one in control over her personal information. In the policies the social media companies claim that the user's privacy and her user-generated data belongs to, and is controlled by, the user. However, later statements contest this by expressing great restrictions and limitations regarding the user's possibilities to fulfill and execute ownership and control. The policies frequently use wordings as "we may", that implies great uncertainties, which can function as a power mechanism that makes it difficult for the users to make informed choices regarding their privacy. In CPM theory ownership and control of private information is considered important components of privacy and this is something that the policies are trying to convey, yet the policies rather express what CPM theory describes as manipulative co-ownership, where the company has the control over access to and disclosure of private information. By using the word "share" instead of "sell" in their privacy policies, the companies try to make business transactions of information to seem more positive. Also the act of "collection" is portrayed as something that is positive, made to improve the user's use of the service. The social media companies' discourses of privacy imply frictions with the viewpoint of users' and privacy researchers' definition of privacy, where sense of ownership and control of private information, the user-generated data, is vital. The policies of Reddit and Ello constitutes exceptions in some respects and their policies express discursive struggle since they correspond with policies of other social media companies, as Ello and Reddit claim they *actually* care about their users' privacy and do not sell the user-generated data. In conclusion, power in the policies is manifested by uncertainties, the users' lack of control and influence and the social media companies' lack of transparency.

By combining interface analysis with discourse analysis of text, this thesis proposes methodological development that could be fruitful for discursive understanding of online structures and contexts. The combination of the theory of critical discourse analysis and communication privacy management theory, also offers a way to integrate theoretical understandings of language and power with theoretical perspectives of privacy.

7.2 Further Research

A great deal of research regarding users' attitudes toward privacy issues has been done, but still there are aspects that call for further investigation. For example, as mentioned in the background, a large amount of the women asked in the survey *Svenskarna och internet 2014* (Findahl 2014) claimed not being able to answer questions about privacy protective measures due to not knowing what the questions referred to. This calls for further studies on privacy awareness and knowledge of Internet technologies, and additionally intersectional perspectives on digital literacy differences are needed, in order to address digital gaps. Also, in the beginning of this thesis I mention that 16 percent of the Swedes believe that privacy is something that no longer exists (Findahl 2014) – is this fight to maintain privacy something that most of us are willing to give up? I believe that this is something that future research needs to confront. Important questions regarding privacy, user-generated data and new technologies concern how public institutions and organizations like libraries, that represent democratic values and has privacy protection written into its ethic codes, should tackle these issues both for the own organization's information management but also regarding the development in society as a whole. Should libraries, next to or as a part of media and information literacy, support privacy literacy to a greater extent? And if so, how should this be accomplished? Further research is needed on whether online privacy issues of user-generated data can be considered a mission for libraries, how these types of issues fit into their democratic mission and what the attitudes regarding this are. What are the viewpoints of librarians? Are they equipped to meet this challenge? It would also be interesting with further knowledge regarding if the library could be a safe space online, like a safe hub to reach policies or other materials, and how this could be designed. Intertwined with these factors are vital questions that need to be discussed; who should stand up for the citizens' privacy? Should libraries educate and publicly speak up concerning these issues? Or should we stand idly by? Last, but in no way least, power structures of information related issues need greater and continuous attention. The power discrepancies between the "data-classes" of the emerging "big-data society" (Manovich, 2011) require acknowledgement and resistance. The lacking power of the public become vital to address as more of us are spending more of our time online and the connectedness of digital tools grows deeper into our lives.

9. References

9.1 References

Altman, I. (2002). Foreword. In Petronio, S., *Boundaries of Privacy: Dialectics of Disclosure*. Albany, NY, USA: State University of New York Press. pp. xiii-xix.

American Library Association (2015a). Choose Privacy Week. https://chooseprivacyweek.org/our-story/privacy-week/ [2015-04-07].

American Library Association (2015b). Why Libraries? https://chooseprivacyweek.org/our-story/why-libraries/ [2015-04-07].

Andrejevic, M. (2015). Personal Data: Blind Spot of the "Affective Law of Value"? *The Information Society: An International Journal*, 31(1), pp. 5-12.

Bechmann, A. (2014). Non-informed Consent Cultures: Privacy Policies and App Contracts on Facebook. *Journal of Media Business Studies,* 11(1), pp. 21-38.

Bergström, G. & Boréus, K. (2008). Diskursanalys. In Bergström, G. & Boréus, K. (Eds.), *Textens mening och makt: metodbok i samhällsvetenskaplig text- och diskursanalys*. Lund: Studentlitteratur. pp. 305-362.

Boréus, K. (2013a). Texter i vardag och samhälle. In Ahrne, G. & Svensson, P. (Eds.), *Handbok i kvalitativa metoder*. Malmö: Liber. pp. 131-149.

Boréus, K. (2013b). Diskursanalys. In Ahrne, G. & Svensson, P. (Eds.), *Handbok i kvalitativa metoder*. Malmö: Liber. pp. 150-164.

Boréus, K., & Bergström, G. (2008). Lingvistisk textanalys. In G. Bergström & K. Boréus (Eds.), *Textens mening och makt: metodbok i samhällsvetenskaplig text- och diskursanalys*. Lund: Studentlitteratur. pp. 263-304.

boyd, d. & Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), pp. 662-679.

Buchanan, M. A. (2011). *Privacy and Power in Social Space: Facebook*. Diss. Stirling: University of Stirling.

Buckland, M. (1991). *Information and Information Systems*. New York: Greenwood.

Burkell, J. & Carey, R. (2011). Personal Information and the Public Library: Compliance with Fair Information Practice Principles *The Canadian Journal of Information and Library Science*, 35(1), pp. 1-15.

Croll, A. (2012). Big data is our generation's civil rights issue, and we don't know it. *Radar*, http://radar.oreilly.com/2012/2008/big-data-is-our-generations-civil-rights-issue-and-we-dont-know-it.html [2015-01-10].

Debatin, B. (2011). Etics, Privacy, and Self-Restraint in Social Networking. In Trepte, S. & Reinecke, L. (Eds.), *Privacy Online: Perspectives on Privacy and Self-Disclosure in the Social Web*. Berlin, Heidelberg: Springer. pp. 47-60.

DeMers, J. (2015). The Top 7 Social Media Marketing Trends That Will Dominate 2015. *Forbes*, http://www.forbes.com/sites/jaysondemers/2014/2011/2019/the-top-2017-social-media-marketing-trends-that-will-dominate-2015/ [2015-03-12].

Dourish, P. & Anderson, K. (2006). Collective Inofrmation Practice: Exploring Privacy and Security as Social and Cultural Phenomena. *Human-Computer Interaction,* 21, pp. 319-342.

Ello (2015d). About. https://ello.co/wtf/post/about-ello [2015-04-05].

Facebook (2015d). Rapport om myndighetsförfrågningar. https://govtrequests.facebook.com/ [2015-01-09].

Facebook (2015e). Company Info. http://newsroom.fb.com/company-info/ [2015-04-02].

Fairclough, N. (2001). *Language and power*. Harlow: Longman.

Fairclough, N. (2013). Critical discourse analysis and critical policy studies. *Critical Policy Studies*, 7(2), pp. 177-197.

Findahl, O. (2014). *Svenskarna och internet 2014*. Stockholm: .SE https://www.iis.se/docs/SOI2014.pdf.

Gibbs, S. (2015). Class action privacy lawsuit filed against Facebook in Austria. *The Guardian*, http://www.theguardian.com/technology/2015/apr/2009/class-action-privacy-lawsuit-filed-against-facebook-in-austria [2015-04-12].

Google (2015f). Insynsrapport. http://www.google.com/transparencyreport/userdatarequests/countries/ [2015-01-09].

Google (2015g). Company. http://www.google.com/about/company/ [2015-04-07].

Google (2015h). Products. http://www.google.com/intl/sv/about/products/ [2015-04-07].

Gressel, M. (2014). Are Libraries Doing Enough to Safeguard Their Patrons' Digital Privacy? *The Serials Librarian: From the Printed Page to the Digital Age*, 67(2), pp. 137-142.

Haider, J. (2008). *Open Access and Closed Discourses: Constructing Open Access as a "Development" Issue*. Diss: London, City University London.

IFLA (2013). *Ridingng the Waves or Caught in the Tide? - Navigating the Evolving Information Environment*. The Hague, IFLA. http://trends.ifla.org/files/trends/assets/insights-from-the-ifla-trend-report_v3.pdf.

Instagram (2015d). FAQ. https://instagram.com/about/faq/ [2015-04-03].

Jensen, C. & Potts, C. (2003). Privacy Policies Examined: Fair Warning or Fair Game? *GVU Technical Report,* 03-04.

Lane, J. I., Stodden, V., Bender, S. & Nissenbaum, H. (eds.) (2014). *Privacy, big data, and the public good: frameworks for engagement.* New York, NY:: Cambridge University Press.

Lenhart, A. (2015). Teens, Social Media & Technology Overview 2015. *Pew Research Center: Internet, Science & Tech*, http://www.pewinternet.org/2015/2004/2009/teens-social-media-technology-2015/ [2015-03-12].

Limberg, L., Sundin, O. & Talja, S. (2012). Three Theoretical Perspectives on Information Literacy. *HUMAN IT, 11*(2), pp. 93-130.

Linkedin (2015d). About us. https://www.linkedin.com/about-us?trk=hb_ft_about [2015-04-04].

Linkedin (2015e). Company LinkedIn. https://www.linkedin.com/company/linkedin [2015-04-04].

Manovich, L. (2011). Trending: The Promises and the Challenges of Big Social Data. In M. K. Golk (Ed.), *Debates in the Digital Humanities*. Minneapolis: The University of Minnesota Press.

Margulis, S. T. (2011). Three Theories of Privacy: An Overview. In Trepte, S. & Reinecke, L. (Eds.), *Privacy Online : Perspectives on Privacy and Self-Disclosure in the Social Web*. Berlin, Heidelberg: Springer. pp. 9-18.

Nina, Ñ. & Boers, R. (2013). Disliking the like: User policy-change and perception of the internet as a democratic medium. *New Library World*, 114(7/8), pp. 319-325.

Nissenbaum, H. & Barocas, S. (2014). Big Data's End Run around Anonymity and Consent. In Lane, J. I., Stodden, V., Bender, S. & Nissenbaum, H. (Eds.), *Privacy, Big Data, and the Public Good: Frameworks for Engagement*. New York, NY: Cambridge University Press. pp. 44-75.

Olson, H. A. (2002). *The Power to Name: Locating the subject representation in libraries*. Boston, Mass: Kluwer.

Olsson, M. R. (2010). Michel Foucault: Discourse, Power/Knowledge, and the Battle for Truth. In G. J. Leckie, Given, Lisa M., Buschman, John E. (Ed.), *Critical Theory for Library and Information Science: Exploring the Social from across the Disciplines*. Santa Barbara, CA: Libraries Unlimited. pp. 63-74.

Petronio, S. (2002). *Boundaries of Privacy: Dialectics of Disclosure*. Albany, NY, USA: State University of New York Press.

Pinterest (2015d). Press. https://about.pinterest.com/en/press [2014-04-03].

Pinterest (2015e). About. https://about.pinterest.com/en [2015-04-05].

Privacy (2010) Oxford dictionary of English. (3 ed.). New York, NY: Oxford University Press.

Reddit (2015e). About reddit. https://www.reddit.com/about/ [2015-04-05].

Reddit (2015f). Frequently Asked Questions. https://www.reddit.com/wiki/faq [2015-04-07].

Shontell, A. (2015). Snapchat Is A Lot Bigger Than People Realize And It Could Be Nearing 200 Million Active Users. *Business Insider UK*. http://uk.businessinsider.com/snapchats-monthly-active-users-may-be-nearing-200-million-2014-12?r=US [2015-04-04].

Snapchat (2015d). Product Evolution. https://support.snapchat.com/a/product-evolution [2015-04-04].

Stanfill, M. (2014). The interface as discourse: The production of norms through web design. *New Media & Society*, pp. 1-16.

State of California (2003). *The California Online Privacy Protection Act of 2003. Internet Privacy Requirements 22575 - 22579*. http://leginfo.legislature.ca.gov/faces/codes_displayText.xhtml?lawCode=BPC&division=8.&title=&part=&chapter=22.&article= [2015-03-12].

Stutzman, F., Capra, R. & Thompson, J. (2010). Factors mediating disclosure in social network sites. *Computers in Human Behavior*, 27(1), pp. 590-598.

Stutzman, F., Gross, R. & Acquisti, A. (2012). Silent Listeners: The Evolution of Privacy and Disclosure on Facebook. *Journal of Privacy and Confidentiality*, 4(2), pp. 7-41.

Stutzman, F. & Kramer-Duffield, J. (2010). Friends only: Examining a Privacy-Enhancing Behavior in Facebook. *CHI 2010 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 1553-1562.

Sundin, O. & Rivano Eckerdal, J. (2014). Inledning: Från informationskompetens till medie- och informationskunnighet. In O. Sundin & J. Rivano Eckerdal (Eds.), *Medie- och informationskunnighet i en biblioteks- och informationsvetenskaplig belysning*. Stockholm: Svensk biblioteksförening. pp. 9-25.

Svensk biblioteksförening (2014). *Barbara Jones talar om intellektuell frihet på Biblioteksdagarna*. Svensk biblioteksförening. http://www.biblioteksforeningen.org/2014/02/20/barbara-jones-talar-pa-biblioteksdagarna/ [2015-02-06].

Trepte, S. & Reinecke, L. (2011). *Privacy Online : Perspectives on Privacy and Self-Disclosure in the Social Web*. Berlin, Heidelberg: Springer.

Trottier, D. & Fuchs, C. (2015). Theorising Social Media, Politics and the State: An Introduction. In D. Trottier & C. e. Fuchs (Eds.), *Social media, politics and the state: protests, revolutions, riots, crime and policing in the age of Facebook, Twitter and YouTube.* New York: Routledge. pp. 3-38.

Tumblr (2015d). Press Information. https://www.tumblr.com/press [2015-04-02].

Twitter (2015d). About. https://about.twitter.com/company [2015-04-05].

United Nations (1948). The Universal Declaration of Human Rights. http://www.un.org/en/documents/udhr/index.shtml#a12 [2015-04-02].

Vaidhyanathan, S. & Bulock, C. (2014). Knowledge and Dignity in the Era of "Big Data". *The Serials Librarian: From the Printed Page to the Digital Age*, 66(1-4), pp. 49-64.

Vasalou, A., Joinson, A. & Houghton, D. (2014). Privacy as a fuzzy concept: A new conceptualization of privacy for practitioners. *Journal of the Association for Information Science and Technology,* 66(5), pp. 918-929.

Wallström, M. (2014). EU pumpar in miljarder i big data. *Computer Sweden*, http://computersweden.idg.se/2.2683/2681.589037/eu-pumpar-in-miljarder-i-big-data [2015-01-02].

Williams, T. L., Agarwal, N. & Wigand, R. T. (2014). Protecting Private Information: Current Attitudes Concerning Privacy Policies. *2014 ASE BigData/SocialInformatics/PASSAT/BioMedCom Conference, Harvard University, (ASE)*, pp. 1-13.

Youtube (2015b). About YouTube. https://www.youtube.com/yt/about/ [2015-04-02].

Youtube (2015c). Statistics. https://www.youtube.com/yt/press/statistics.html [2015-04-04].

Zimmer, M. (2013). Assessing the Treatment of Patron Privacy in Library 2.0 Literature. *Information Technology and Libraries*, 32(2), pp. 29-41.

9.2 Empirical Material

Ello (2015a). Ello. https://ello.co/ [2015-03-10].

Ello (2015b). Ello Policies. https://ello.co/wtf/post/policies [2015-03-10].

Ello (2015c). Ello Privacy Policy. https://ello.co/wtf/post/privacy [2015-03-11].

Facebook (2015a). Facebook. https://www.facebook.com [2015-03-10].

Facebook (2015c). Data Policy. https://www.facebook.com/policy.php [2015-03-11].

Facebook Inc (2015b). Facebook (Version 26.0). [Mobile application]. https://itunes.apple.com/app/facebook/id284882215.

Google (2015a). Google. https://www.google.com/ [2015-03-10].

Google (2015d). Privacy Policy. http://www.google.com/policies/privacy/ [2015-03-11].

Google (2015e). Privacy & Terms: Key terms. http://www.google.com/policies/privacy/key-terms/ - toc-terms-info [2015-03-22].

Google Inc (2015b). Google (Version 5.2.43972). [Mobile application]. https://itunes.apple.com/se/app/google/id284815942?mt=8.

Google Inc (2015c). Youtube (Version 10.09.11538). [Mobile application]. https://itunes.apple.com/se/app/youtube/id544007664?mt=8.

Instagram (2015a). Instagram. https://instagram.com [2015-03-10].

Instagram (2015c). Privacy Policy. https://instagram.com/about/legal/privacy/. [2015-03-10].

Instagram Inc (2015b). Instagram (Version 6.7.2). [Mobile application]. https://itunes.apple.com/us/app/instagram/id389801252?mt=8&ign-mpt=uo%3D2.

Linkedin (2015a). Linkedin. https://www.linkedin.com [2015-03-10].

Linkedin (2015c). Your Privacy Matters. https://www.linkedin.com/legal/privacy-policy?trk=hb_ft_priv [2015-03-11].

LinkedIn Corporation (2015b). LinkedIn (Version 8.6). [Mobile application]. https://itunes.apple.com/se/app/linkedin/id288429040?mt=8.

Pinterest (2015a). Pinterest. https://www.pinterest.com [2015-03-10].

Pinterest (2015c). Privacy Policy. https://about.pinterest.com/sv/privacy-policy [2015-03-11].

Pinterest Inc. (2015b). Pinterest (Version 4.4.1). [Mobile application]. https://itunes.apple.com/se/app/pinterest/id429047995?mt=8.

Reddit (2015a). Reddit. http://www.reddit.com [2015-03-10].

Reddit (2015b). Wiki. http://www.reddit.com/wiki/index [2015-03-10].

Reddit (2015c). Alien Blue - reddit official client (version 2.9.2). [Mobile application]. https://itunes.apple.com/se/app/alien-blue-reddit-official/id923187241?mt=8.

Reddit (2015d). Reddit Privacy Policy. http://www.reddit.com/wiki/privacypolicy [2015-03-10].

Snapchat (2015a). Snapchat. https://www.snapchat.com [2015-03-10].

Snapchat (2015c). Privacy Policy. https://www.snapchat.com/privacy [2015-03-10].

Snapchat Inc. (2015b). Snapchat (version 9.3.0). [Mobile application]. https://itunes.apple.com/se/app/snapchat/id447188370?mt=8.

Tumblr (2015a). Tumblr. https://www.tumblr.com [2015-03-10].

Tumblr (2015b). Tumblr (version 3.8.2). [Mobile application]. https://itunes.apple.com/se/app/tumblr/id305343404?mt=8.

Tumblr (2015c). Privacy Policy. https://www.tumblr.com/policy/en/privacy [2015-03-11].

Twitter (2015a). Twitter. https://twitter.com [2015-03-10].

Twitter (2015c). Twitter Privacy Policy. https://twitter.com/privacy?lang=en [2015-03-10].

Twitter Inc. (2015b). Twitter (version 6.23.1). [Mobile application]. https://itunes.apple.com/us/app/twitter/id333903271?mt=8&ign-mpt=uo%3D2.

Youtube (2015a). Youtube. https://www.youtube.com [2015-03-10].