RUHR
UNIVERSITÄT
BOCHUM

RUB

RUHR-UNIVERSITÄT BOCHUM

Digital Oblivion in Online Social Networks: A Necessity or Just Nice to Have?

Maria Kober

Master's Thesis – October 21, 2019.
Mobile Security Group

Supervisor: Prof. Dr. Markus Dürmuth
Advisor: Florian Farke, M.Sc.

hgi MobSec

Abstract

Digital technology makes forgetting difficult. The term *digital oblivion* summarizes the transfer of forgetting to the digital world. The first contribution of this thesis is an overview of arguments for and against digital oblivion found in literature. The debate of whether digital oblivion should be implemented or not is controversial. Both sides state that the absence or presence of forgetting mechanisms introduces censorship, restricts the freedom of speech, and presents a danger to democracy. The second contribution of this thesis is to answer the question of whether the absence of forgetting mechanisms in online social networks (OSN) is a problem for users. This question was answered by conducting a user study with 250 participants. Users would appreciate tools implementing digital oblivion in OSN to take action against data that is spread about them against their will, to check if their content is offline after they deleted their account, and as an optional feature to automatically delete their content after a fixed time. Users do not want their content to be automatically deleted. The presence of tools implementing several facets of digital oblivion would be appreciated and considered helpful by users of OSN.

Eidesstattliche Erklärung

Ich erkläre, dass ich keine Arbeit in gleicher oder ähnlicher Fassung bereits für eine andere Prüfung an der Ruhr-Universität Bochum oder einer anderen Hochschule eingereicht habe.

Ich versichere, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen benutzt habe. Die Stellen, die anderen Quellen dem Wortlaut oder dem Sinn nach entnommen sind, habe ich unter Angabe der Quellen kenntlich gemacht. Dies gilt sinngemäß auch für verwendete Zeichnungen, Skizzen, bildliche Darstellungen und dergleichen.

Ich versichere auch, dass die von mir eingereichte schriftliche Version mit der digitalen Version übereinstimmt. Ich erkläre mich damit einverstanden, dass die digitale Version dieser Arbeit zwecks Plagiatsprüfung verwendet wird.

Official Declaration

Hereby I declare that I have not submitted this thesis in this or similar form to any other examination at the Ruhr-Universität Bochum or any other institution or university.

I officially ensure that this paper has been written solely on my own. I herewith officially ensure that I have not used any other sources but those stated by me. Any and every parts of the text which constitute quotes in original wording or in its essence have been explicitly referred by me by using official marking and proper quotation. This is also valid for used drafts, pictures and similar formats.

I also officially ensure that the printed version as submitted by me fully confirms with my digital version. I agree that the digital version will be used to subject the paper to plagiarism examination.

Not this English translation but only the official version in German is legally binding.

Datum / Date	Unterschrift / Signature

Contents

1 Introduction

Forgetting is essential for individuals as well as for societies. The ability to forget lets people act and live in the present without being restrained by the past [1]. Forgetting on a societal level enables an individual or a group of individuals to get past mistakes or wrongdoings, giving them the freedom to reinvent and improve themselves [2].

Besides forgetting, remembering is also of particular importance for individuals and societies [2]. As Mayer-Schönberger [3] points out, remembering helps to avoid making costly and dangerous mistakes twice.

History shows that humans constantly strove to develop strategies to increase their ability to remember important events and share knowledge [1, 3]. Many stories and epics were passed on to following generations as oral tradition. Techniques and methods have been invented to serve as external memories – paintings on walls or canvas, scripture and pictograms used in books, photography, and video technology [3].

The invention of computers and the subsequent emergence of modern information technology changed how humans can externalize their memory. Instead of needing paper for text, canvas for paintings, and tape for audio records, it is possible to store all of these contents on a single device represented by digital bits.

In the last thirty years, computer memory space and easy-to-use recording tools like cameras have become mass-market products [3–5]. This technology allows individuals to produce and store Terabytes of data [3, 6]. Techniques for digital information retrieval have been developed and refined, making it easy and fast to search through stored data [3]. Thus, computer technology is a way to support remembering for individuals and society as a whole [1].

People do not only store personal data on devices at home. In 2017, about $2.5 \cdot 10^{18}$ bytes of data were created per day and uploaded to the world wide web, often publicly or semi-publicly accessible to more than one person [3, 7]. Users of online services create a considerable amount of content on the world wide web; e. g., in 2017, two-thirds of internet users were active users of online social networks (OSN) [8] and in 2018, more than 600.000 posts were created on the social media platforms

Twitter[1], Tumblr[2], and Instagram[3] every minute [9]. However, there is also data people do not upload deliberately but the services they use collect and store about them. This automatic collection of data may include, for example, search engine queries, travel dates and preferences, and other online behavior. A human would forget many of these queries as well as uploads on OSN, but the data still exists years after the event [3].

Mayer-Schönberger [3] states that today, "forgetting has become the exception, and remembering the default." Bannon [1] points out that "there has been very little consideration of the use of technologies to help us forget, either at an individual level or at a group or society level."

The term *digital oblivion* summarizes the transfer of forgetting to the digital world. This transfer is not a single step or instruction but has many facets like public discussion, jurisdiction, proposals of implementation, or contributions of several scientific research fields.

1.1 Contribution

The contribution of this thesis is two-fold.

First, this thesis presents an overview on arguments for and against digital oblivion found in literature. This overview gives readers the opportunity to understand both sides of the discussion on whether digital oblivion should be implemented or not.

Second, this thesis contributes answers to the question whether the absence of forgetting mechanisms in OSN is a problem for users themselves. This question is answered by conducting a user study. The results indicate that users would appreciate forgetting mechanisms for most facets of digital oblivion.

1.2 Related Work

Novotny et al. [10] conducted a user study to investigate "oblivion on the web". They found that users wish for oblivion due to three reasons: privacy concerns over disclosed private information, having control over their data, and wanting to dissociate from data the users perceive as obsolete.

[1] `https://www.twitter.com`
[2] `https://www.tumblr.com`
[3] `https://www.instagram.com`

Time Dimension of Data Sharing and Data Deletion

Several studies investigate the question whether deletion behavior and sharing preferences of content in OSN change when content ages.

Bauer et al. [11] investigated how privacy preferences of Facebook[4] posts change over time. They found that users do not want content to be entirely deleted when it gets older. Instead, users want some content to become more private and some content to become more visible.

In their study, Murillo et al. [12] came to similar results. They found that users do not want their content to be deleted automatically. Instead, the value of data for users changes over time; sometimes, the value of information increases through specific events. Additionally, the value of data depends on the context where the information has been published in.

Mondal et al. [13] investigated deletion behavior on Twitter. They found that users intentionally delete recent tweets more often than they delete older tweets. Old tweets are more often deleted through account deletion than active deletion by the user who had published it.

Ayalon et al. [14] conducted a study on sharing preferences on Facebook. They found that the relevance of a post and the willingness to share a post decrease with time. Several users wanted to delete their posts when they deem the posts irrelevant due to being old.

In another study, Ayalon et al. [15] came to similar conclusions. They found that the willingness to share posts on Facebook decreases with time. They did not find any indication that the majority of users wants to alter or delete their old posts.

Deletion, Content Distributed by Others, and Trust in Social Network Providers

Several studies contribute to a better understanding of how users perceive the deletion of content that has been published online. Other studies contribute to the question how users manage information that is spread by others in OSN.

Murillo et al. [12] investigated how users understand the deletion of data, including the deletion of data that has been published on social media. They found that the main reasons for deletion in social media are because the data is considered outdated or the content is potentially embarrassing. Regarding data storage, about half of their participants believed that some of their deleted data stays on the social network provider's servers or that some data can not be removed at all.

In their work, Thomas et al. [16] analyzed privacy issues related to Facebook friends. They point out that maintaining control over information about oneself in social

[4] `https://www.facebook.com`

media is almost impossible. Any user can share information about another user without the latter having any control over it or even knowing about it. Lam et al. [17] conducted a study on personal information disclosure in OSN through others. They came to a similar conclusion: it is impossible to control how one's own personal information is disclosed.

Besmer et al. [18] investigated sharing and tagging of pictures on Facebook. They found that users are concerned about not being able to remove images from Facebook that were posted by others. Their participants fear that those images might have an influence on them in the future. To take pictures down, many of their participants rely on offline interaction like talking to their friends. This strategy did not always result in an image being taken down.

Madden [19] found that more than 40% of users in OSN deleted comments that others had made on their profiles. She also found that 37% of users remove their names from photos where they had been tagged.

Madden et al. [20] report that almost 10 % of the internet users in the United States of America had asked someone to remove information about them that was posted online. They found that 82% of those requests were successful, yet they do not describe how users asked for the information to be taken down. According to their study, the majority of users do not trust social network service providers.

In their study, Lampinen et al. [21] investigated how users manage disclosure of information by others in OSN. Where possible, their participants remove tags linking them to the information or remove comments on their profiles. In case another user is in control of the information, they ask him to remove it. If this request fails, they report inappropriate content to the service provider and ask for removal.

Automated Data Expiration

Several authors contributed proposals for automated data expiration.

Perlman [22] introduced *The Ephemerizer*. This tool is based on cryptographic keys created, stored, and managed by a central key management service. Data that is meant to expire is encrypted with those keys. As soon as the expiration date is reached, the key is destroyed and the data becomes unreadable.

Geambasu et al. [23] introduced the system *Vanish*. This system utilized cryptographic techniques to ensure that all copies of certain data becomes unreadable after a set expiration date. *SafeVanish* [24] is an improved version of *Vanish*.

Backes et al. [25] developed the system *X-pire!*. They use keys stored on a centralized keyserver to add an expiration date to images shared online. The image becomes unavailable once the expiration date is reached.

Castelluccia et al. [26] introduced the *EphPub* protocol to prevent access to expired content. This protocol is based on encryption and the Domain Name System (DNS).

Reimann et al. [27] introduced a data revocation system that utilizes that websites constantly change over time.

Zarras et al. [28] developed a protocol for data expiration based on cryptographic keys and the Domain Name System (DNS). This protocol does not use a given expiration date, but uses heuristics to expire data access once the interest in the data decreases.

The messaging service Snapchat[5] automatically deletes messages from their servers after a set time, usually 24 hours after the message has been read. After this time, the messages can not be viewed by receiver or sender [29].

1.3 Outline

The second chapter presents arguments for and against digital oblivion found in literature. Additionally, it presents examples that are often used to motivate digital oblivion and introduces a definition of digital oblivion in the context of online social networks.

The third chapter introduces the research question, describes the study design and how the participants were chosen.

The fourth chapter presents the results of the survey.

The fifth chapter interprets and discusses the results of the study and points out limitations.

The sixth chapter provides a summary and points out future work.

[5] `https://www.snapchat.com`

2 Background

The first part of this chapter introduces arguments for and against digital oblivion found in literature. It includes events or incidents motivating the need for digital oblivion.

The second part of this chapter introduces the definition of digital oblivion in the context of online social networks (OSN) as it is understood in the further course of this thesis.

2.1 Arguments Regarding Digital Oblivion in Literature

This section introduces arguments for and against digital oblivion found in literature. It starts by presenting arguments for digital oblivion, followed by examples used to motivate the necessity of digital oblivion. The last part of this section presents arguments against digital oblivion.

2.1.1 Arguments for Digital Oblivion

In several papers and books, authors provide arguments implying that digital oblivion is necessary. Some argue for forgetting mechanisms on a general base and point out the benefits of such mechanisms. Other authors warn of consequences in case forgetting mechanisms are not introduced to technology.

Loss of Information Sovereignty

Several authors expressed their concern that people are losing control over information which is shared by them or about them on the Internet. Thus, people are inevitably losing their information autonomy, including control over their privacy and their personal information [30–32]. Conley [33] warns that "without control over our own information, we are vulnerable to external forces – and this vulnerability affects the way we think, behave, and grow."

Self-Published Content
Someone who publishes data on the Internet loses control of this data forever as soon as it is online. In case the original source of data is deleted, the data might still

remain somewhere online, e. g., as archive or screenshot. The person publishing the data in the first place has no say whether or when this information will be forgotten by the public [3, 34, 35].

Mayer-Schönberger [3] points out that most users are not fully aware of which information is known about them. The majority of users are not aware of risks connected to the sharing of information online like, for example, how this public information might shape their future.

Most users are not aware of the level of exposure they are facing when using digital online services. Additionally, the majority of users in OSN does not change the sharing settings of content they posted in the past. Old settings remain even if those settings are out-dated or not in accordance with the current visibility preferences of the user [3, 36].

Information Disclosed by Others
Several authors point out that the publisher of a piece of information might not be the person the information is about. With today's easy access to the Internet and widespread mobile recording technology, anyone can take a picture of another person and publish it online. The distribution of information, such as images, is out of control of the individual it is about. A person can be linked to content through mechanisms like photo tagging, making it even harder to control who accesses this information [3, 37].

Gossip and rumors can be spread easily by friends, family members, co-workers, or adversaries. Not only that but (family) secrets or the content of personal emails can be shared with a broad public as well [37, 38].

Data Collected by Third Parties
People lose control over who accesses and collects their data. Search engines, web services, and device sensors all collect data, including search history, location data, browsing habits, and reading behavior. Users do not have control over the data which is collected or disclosed through those services. Users might not even know that this data, including personal data, is collected about them. Search engines and web services store and remember data that the user might long have forgotten [3, 39, 40]. Kieselmann et al. [34] state that "[i]nternet services existing today entail that people knowingly or unknowingly share more and more personal data."

Online services and databases organize information available online and make it easy to retrieve, analyze, and utilize this data. Technology allows for various forms of data matching, de-anonymization, and datamining, resulting in extensive digital dossiers [39–41]. Haga [42] states that "[n]owadays [...], [r]ealistically, no one can know the entire processing of personal information on the Internet." Ambrose et al. [39] point out that "the few efforts people can make in order to 'protect their privacy' online, are often ignored or circumvented."

Data Access Through Third Parties
A person cannot control who accesses information about them. For example, employers might check the online reputation of people they intend to hire. False information shared by others about a person online can lead to problems for this person when applying for a job [39, 43]. Bishop et al. [2] warn that it will be easy for businesses, adversaries, and curious people to "discover unpleasant truths about individuals."

Conley [33] warns that "[t]he preservation of even innocuous information can have disturbing consequences when that information is aggregated." He states that companies and governments learn a lot about people's private lives through their public online activities.

Mayer-Schönberger [3] points out that information might be given to (third) parties the user is not aware of. As an example, he points out that about two-thirds of U.S. health insurance companies access the digital prescription histories of health insurance applicants.

No Secure Mechanisms for Deletion or Correction
Several authors point out that information that has been shared online might stay online, regardless of the steps someone takes to get this information removed. Likewise, it is difficult to correct information on the Internet [31].

Even when content is deleted, it might have been downloaded, copied, shared, reposted, duplicated, mirrored, cached by search engines or archived elsewhere online since the day it has been posted deliberately or accidentally. This makes it impossible to say whether deleted content is gone from the Internet for good [25, 30, 32, 34]. Kieselmann et al. [34] point out that content can still be found years after it has been first published, even if the original source has already been deleted. Some authors go as far as to say that information once published online is "available essentially forever" [25] and "preserved for eternity" [2].

Novotny et al. [10] point out that "[d]ata brokers link the many traces people leave online and preserve duplicates in databases for eternity."

Digital Baggage and Ruined Reputation

Some authors point out that, with digital technology, remembering instead of forgetting becomes the norm. Mayer-Schönberger [3] states that "we have begun to unlearn forgetting" and Haga [42] points out that "with the Internet and digital technology, the world has forgotten how to forget."

Permanent Digital Baggage and Paralysis
A variety of information – e. g., texts, images, and videos – is stored online. This information is only one click away from access by the public. It can become permanent digital baggage and a burden for a person [37, 42]. Rosen [44] notes that "far from

giving us a new sense of control over the face we present to the world, the Internet is shackling us to everything that we have ever said, or that anyone has said about us, making the possibility of digital self-reinvention seem like an ideal from a distant era."

Humans usually forget or change the memory of what happened in the past over time. Having perfect digital memory at hand might lead people to distrust their own memory of past events. Also, people might start to overvalue details of their lives which they would usually forget quickly [3, 45]. Mayer-Schönberger [3] points out that "[p]erfect remembering exposes us to filtering, selection, and interpretation challenges that forgetting has mostly shielded us from."

Too much available information about individuals has the potential to "stifle and control people's lives, rather than act as a liberatory force" [1], to induce "a state of paralysis, affecting people's ability to act" [1], and might lead to people losing their ability to "live and act firmly in the present" [3]. People might become self-absorbed as they try to create a sufficient self-presentation all the time [45]. Bannon [1] expands previous statements to organizations, pointing out that "computer archives can become the source of organizational paralysis, stifling innovation and creativity by channelling people's efforts into examining only the accumulated material of the past, rather than spending time on exploring possible futures."

Haga [42] warns that a lot of data is created without knowledge of the person it is about. This information may become obsolete over time, yet at any time, it has the potential to cause damage to the individuals it is about and to become a serious obstacle in their lives.

Non-Forgiveness and Ruined Reputation

Loss of information control might lead to a future where nothing is forgiven because nothing is forgotten. Content posted online may influence a person's life years afterwards, even when the event this content is about has been long forgotten by the human mind [3, 33, 43]. For example, Kieselmann et al. [34] and Andrade [46] point out that information found online may influence a person's chance when applying for a job.

Ambrose [47] states that "old information threatens harsh and wide-reaching consequences to the socially valued and often protected individual interests of reputation, identity, and rehabilitation." Life and reputation of individuals might be permanently ruined by the information that can be found online. This information might be about momentary wrongs or misdemeanors – things that would be quickly forgotten without technology. It could be gossip or shaming by others, made public and permanent online. This kind of public information not only damages a person's reputation, but it also rids people of the ability to be who they want to be [37, 42, 48].

Ambrose et al. [39] point out that "in today's information society, it is practically impossible to predict (all) negative consequences of the use of personal data." Infor-

mation from the past might be damaging for a person in the present or the future. Details of a person's life might be presented outside of the context they occurred in or overlay a person's public identity [32, 46]. Bishop et al. [2] state that "[i]n the digital age [...] we are not afforded the luxury of burying past mistakes and starting over. Those mistakes are always one Google search away, disallowing a return to obscurity."

Evolving Through Forgetting

Several authors write about the advantages of societal forgetting. Mayer-Schönberger [3] points out that "societal forgetting gives individuals who have failed a second chance." Forgetting gives people the chance to be different from their past selves, to evolve and improve themselves over time, to reinvent themselves, and to interact with others without constant reminders of what they did in the past [2, 3, 46].

Mayer-Schönberger [3] warns that the combination of information being widely available to others and loss of information control constricts "the freedom to shape one's own identity" and the ability to define oneself.

Freedom of Expression and Self-Censorship

Information about a person can be remembered and made publicly available for longer than the person's lifetime [3]. Several authors express their concern that this circumstance might lead people to stop to express themselves freely or that people might censor what they say.

Zittrain [48] warns that people might "moderate themselves instead of expressing their true opinions." Korenhof et al. [49] warn that if nothing can ever be forgotten, debates may "be stifled or curbed for fear of future consequences later on in life." Mayer-Schönberger [3] sees the danger of self-censorship when all a person expresses could be re-interpreted later. In their study, Sleeper et al. [50] found that people self-censor content they post on Facebook. The reason for this self-censorship is that the content which users post might be seen by other audiences than originally intended.

Bannon [1] takes those previous concerns a step further, arguing that archiving everything on computers "can lead to a stultification in thinking." He argues that people might be afraid to act. They might fear how their doings will be interpreted in the decades to come or which information of their past might be found and re-interpreted. As an example, he points to politicians, "who [...] find aspects of their past being used to criticize them." Bishop et al. [2] are concerned about everyday consequences, stating that "[i]f individuals fear reprisals for lawful but embarrassing conduct, they may feel overly constrained and stressed." Mayer-Schönberger [3] points out that this could be a danger to democracy as it might keep people from protesting against corporations or governments.

Bannon [1] is concerned that human interaction might be reduced. As reason, he points out that technology enables people to record and archive all that is said or done in a moment. Previous to the existence of such recording technology, this information would only be available to those who are present while the interaction takes place. Similarly, Blanchette et al. [51] are concerned that humans change their behavior. They fear that people become different personalities than they would become without recording technology. They argue that "the mere fact that one is being watched changes the way one behaves" and that individuals "come to see themselves as they believe they are seen by their watcher."

False Representation of a Person

Information which is found online about a person might not represent this person accurately.

People's interests, skills, opinions, and views change over time. Information found about individuals online might not represent their current views and opinions. This is especially true for user profiles in OSN. Information about different time periods accumulates there, possibly leading to a false representation of who the person is today. Important events or pieces of information might not have been recorded and thus are not available when viewing a person's profile. This is another factor contributing to an incomplete or false representation of a person [3, 34, 52].

When information ages, chances increase that this information is irrelevant, inaccurate, or presented outside of the original context. The original context of the information might even get lost over time, thus increasing the chance that the information is not truthful anymore [3, 46, 47].

Information tends to disappear over time. Negative and harmful information tends to remain available for longer than other content. Such data can become detrimental misinformation about a person [47].

Information can be interpreted differently by different people and in different time periods [3]. As an example, Mayer-Schönberger [3] refers to medical information collected over time. The interpretation of records and symptoms can vary over time and depend on the person interpreting them.

Data sets belonging to different people are compared for similarities and differences. Based on these computations, a person's profile is classified and judged. This classification might be erroneous and not represent a person well [3]. Mayer-Schönberger [3] calls this "a digital and much broadened version of guilt by association." Kwak et al. [31] point out that online reputation systems are often biased as accurate ratings and feedback are not available. Incorrect information might be harmful to one's online and offline reputation.

Danger of Unintended or Malicious Data Usage

Information which is helpful in one case can be embarrassing, dangerous, or even life-threatening in other cases, when used in another way than intended. Potentially harmful information might be publicly available or it might be collected and archived by organizations or governments [3, 48]. As an example from the last century, Mayer-Schönberger [3] points to the population registry in the Netherlands in the 1930s, which was intended for welfare programs. After invading the Netherlands, the Nazis used this register to identify and deport Jews and Gypsies.

Data often remains available online after it has lost its relevance. This opens the door to misuse by third parties either out of curiosity or with malicious intent. The same issues arise with non-publicly stored data which can become part of data leaks. Additionally, service providers may intentionally misuse private user data [26, 32, 34].

Information might be disclosed in one context, but not in another. Mayer-Schönberger [3] points out that, with today's technology and synchronization mechanisms, information pieces from different sources can be linked together. Personal data might be used for making inferences about the individual it describes. The individual might not expect those inferences to be made by the service provider or might be surprised to know which data the service provider has access to [3, 10]. Novotny et al. [10] point out that employers use web search engines and social media sites to run background checks on job applicants. They report that recruiters rejected candidates because of the information they found online.

Surveillance Technology

Some warn that technology can be used for surveillance by tracking human activity as comprehensive as possible. Mayer-Schönberger [3] states that extensive digital memory represents a digital panopticon, surveying people "in every corner [...] [and] across time." Castelluccia et al. [26] warn that everlasting information "may become prey of [...] government surveillance." Blanchette et al. [51] point out that citizens living in an environment of constant surveillance present a danger to democracy. They fear that people might stop to think critically or take action.

2.1.2 Examples Motivating Digital Oblivion

This section introduces examples found in literature which are used to motivate and advocate digital oblivion. The examples center around incidents that involve specific individuals.

Several papers do not include names or nicknames of a person. Instead, they hint on an incident, like someone being denied her teaching degree due to an image found online. These papers are not included as reference in this section. They are excluded to avoid wrong association with an incident. General arguments mentioned in literature, like people losing their jobs or having reputation issues due to information found online, are not included in this section.

Stacy Snyder and the Drunken Pirate

The example of Stacy Snyder and her picture titled "Drunken Pirate" is used by various authors to point out negative consequences of non-forgetting in OSN [2, 3, 38, 42, 44, 47, 53].

In 2006, Stacy Snyder was about to become a teacher. She had finished her university coursework as well as practical training. Nonetheless, she was denied her teaching degree because university officials saw her behavior as not fit for a teacher. The reason for this opinion was a photo on MySpace[1] where she was seen wearing a pirate hat and drinking from a plastic cup. The picture was titled "Drunken Pirate", implying that she was drinking alcohol – a legal activity in her spare time. As the image was publicly available on the social media platform, it was interpreted as a potential exposure of pupils to their teacher drinking alcohol. The picture could not be taken down successfully as the page had been cataloged and archived by search engines and web crawlers.

Several articles [54, 55] point out that the denial of Stacy Snyder's teaching degree was caused by academic reasons and not by posting a picture of herself online. A lawsuit was filed against the university [38], so no official comment could be made by university officials when the story started being published [54, 56].

People Losing Their Jobs

Mayer-Schönberger [3] gives the example of sixteen-year-old Kimberly Swann. She lost her job because she mentioned that her job was boring on Facebook [57].

Ambrose [47] gives the example of eighteen-year-old Caitlin Davis. She was fired after images appeared online where she was posing next to a passed-out man. The man's skin was covered with sharpie markings, including a swastika symbol.

[1] `https://www.myspace.com`

Newspaper Articles, Journal Publications, and Encyclopedia Entries

Several authors [31, 42, 58–60] mention the case of Mario Costeja González. In the 1990s, a Spanish newspaper printed a notice showing the foreclosure of his house. In 2010, González wanted this notice to be taken offline. He argued that this information was in no way relevant to his current financial status and that deletion was necessary to maintain his honor. Also, withdrawal of the notice would not harm social interests. Additionally, he wanted this information to disappear from the search results of Google[2] and Google Spain[3]. The Court of Justice of the European Union decided this case. According to the court decision, the newspaper could maintain the contents for the purpose of freedom of speech. Google Spain had to remove their links to the newspaper article. Bunn [60] states that this case introduced the legal debate resulting in the right to be forgotten in the General Data Protection Regulation of the European Union [61].

Mayer-Schönberger [3] gives the example of the attorney Shakespear Feyissa. Whilst being at university, he was charged on suspicion of attempted sexual assault but never arrested. The university student newspaper published this case. Ten years later, when Feyissa had his own company, the student newspaper article was one of the top hits on Google search when searching for Feyissa's name. After several years, Feyissa managed to get the article taken down [62].

Mayer-Schönberger [3] mentions the case of psychotherapist Andrew Feldmar. In 2006, Andrew Feldmar wanted to cross the border from Canada to the U.S.A., like he had done several times before. To his surprise, he was denied entry to the United States of America. The border guard searched for him on an online search engine and found an article where Feldmar mentioned that he had taken LSD in the 1960s. Despite having no criminal record and not taking drugs since 1974, he was denied entry to the U.S.A.

Bishop et al. [2] give the example of Wolfgang Werlé who murdered the actor Walter Sedlmayer in 1990. After being released in 2007, his lawyers demanded the Wikimedia Foundation[4], operators of the online encyclopedia Wikipedia[5], to delete Werlé's name from the Wikipedia article related to the victim. They justified this request by stating that Werlé's rehabilitation and future life outside the prison system were strongly impacted by his name being present in the Wikipedia article [63]. It should be noted that this case is also used to argue against a too general approach to digital oblivion [47, 64]. One of those speaking against the removal of the Werlé's name states his concerns as follows: "[a]t stake is the integrity of history itself. If all publications have to abide by the censorship laws of any and every jurisdiction just because they

[2] `https://www.google.com`
[3] `https://www.google.es`
[4] `https://www.wikimediafoundation.org`
[5] `https://www.wikipedia.org`

are accessible over the global internet, then we will not be able to believe what we read" [63].

Information and Images Disclosed by Others

In their works, Ambrose [47] and Zittrain [48] mention Gyslain Raza as "Star Wars Kid". In 2003, Gyslain Raza made a video where he is playing Star Wars with a golf ball retriever. This video was found by others some time later. They shared the video on the Internet with a broad public without the knowledge or consent of the person shown in the video. For years afterwards, Gyslain Raza was known as "the Star Wars Kid".

Solove [37] and Zittrain [48] give the example of the "dog poop girl". A young woman's dog pooped in a subway train in North Korea. The woman refused to clean it up. Someone took photos of her and posted the incident on a Korean blog. This blog entry started a series of parodies and privacy disclosures about her history and her family. As a result of the public shaming and embarrassment, the young woman dropped out of university and quit her job.

Stokes et al. [65] give the example of Amanda Todd. Amanda Todd had sent an explicit image to a man she met in a chatroom. First, he tried to blackmail her. When blackmailing did not work, the man repeatedly used Facebook to send the image to all her friends. This disclosure resulted in embarrassment and bullying for Amanda. When she created a new Facebook account and changed school, the man sent the image to her new social environment. As a result, Amanda Todd committed suicide at the age of fifteen.

Zittrain [48] mentions the "Bus Uncle of Hongkong" as an example. The "Bus Uncle" upbraided a fellow bus passenger who asked him to speak more quietly on his mobile phone. Another passenger recorded this incident and uploaded it to the Internet where the video was viewed more than a million times. This lead to a number of parodies about the incident and a physical attack on the "Bus Uncle" several weeks later.

2.1.3 Arguments Against Digital Oblivion

Several authors speak out against an implementation of digital oblivion. They warn of negative consequences in case forgetting mechanisms are introduced to technology. Other authors point out the advantages of open data access and data processing by organizations and individuals.

Censorship and Violation of Freedom of Expression and Speech

Most authors voiced their concerns in the debate around or as a response to the *right to be forgotten* in the General Data Protection Regulation of the European Union [61]. Thus, most arguments in this section refer to a specific implementation of digital oblivion.

Freedom of Expression and Speech
Several authors worry that a right to be forgotten might conflict with other rights such as the freedom of expression and the freedom of speech [30, 39]. Some state that the right to be forgotten not only conflicts with those two rights but restricts, contradicts, or threatens them [31, 33, 35, 38, 42, 66].

Rosen [38] states that "[proposals for] new legal rights of oblivion that would allow us to escape our past, these rights pose grave threats to free speech." He names the right to be forgotten "the biggest threat to free speech on the Internet in the coming decade" [64] which could lead "to a far less open Internet" [64].

The freedom of expression is in danger as companies and individuals could easily demand that content created by another individual is taken offline. The freedom of speech is in danger as truthful information can be taken down upon request by the person it is about, even when someone else published the information. Thus a right to be forgotten limits the possibilities of a person to speak about another person [33, 35, 39, 46, 53, 66]. Garcia-Murillo et al. [35] summarize their concerns in a question: "[t]o what extent should individuals have this right to ask others to delete [truthful] information about them that they do not control?".

Right to Know and Limited Access to Information
Deleting information found online results in reduced or limited access to information for individuals [30, 39, 42, 67]. Kent Walker [67] states that the right to be forgotten "represents a serious assault on the public's right to access lawful information."

Haga [42] points out that people would try to delete all inconvenient information that can be found online if there existed an easy possibility for deletion. Steps to take content offline would be made even if there are justifications to have this information available to society. He warns that "[l]arge-scale erasure of data would severely inhibit access to information and the right to know." Kent Walker [67] points out that forced deletion would eliminate the right to know information about representatives, like politicians, or people and organizations providing services.

Haga [42] fears that a right to be forgotten might demotivate or intimidate people to put information online in the first place. He warns that this behavior would "have a dramatic effect, suppressing the free flow of ideas and discourse needed in an open society."

Douglas [59] warns that, in practice, the right to be forgotten will lead to information being removed even when there is doubt about the necessity. Information is precautionary removed to limit the risks and costs of companies.

Censorship
Several authors point out that a right to oblivion can be seen or used as a mechanism for censorship.

Ambrose et al. [39] name the right to be forgotten a "concealed form of censorship." It can be misused to block content or take down truthful information relating to a person. Thus, it restricts the freedom of expression and introduces censorship [30, 64].

Conley [33] notes that such a right may have an impact on the freedom of the press. He notes that it would eliminate the press's ability to (re)publish incidents in case those would become relevant again. Fleischer [66] warns that "more and more, privacy is being used to justify censorship."

Danger to Democracy
Some see a right to oblivion as undermining or endangering democracy. Freedom of expression and freedom speech are necessary for political discourse and thus essential for democracy. Access to information is the foundation of open discussion and discourse. It gives individuals the possibility to gain knowledge and to consider and make informed decisions [35, 37, 42, 59]. Douglas [59] points out that access "to information that is critical of other individuals" is of particular importance when making decisions. Kent Walker [67] warns that a right to erasure would prohibit the public's knowledge of important information about their representatives.

Erasing and Rewriting History

Laws enforcing digital oblivion can be used to rewrite the past and erase aspects of history. This might endanger the integrity of historical records found online [2, 42, 47]. Rosen [38] warns that people might ask for content deletion when they decide to run for a political office or official position. This behavior would restrain people's ability to make informed decisions in democratic processes.

Some authors are concerned about the integrity of long-term historical records. Rosen [38] compares deleting and de-listing of content through a right to be forgotten to the "Orwellian vision of rewriting history on a selective basis." Andrade [46] points out that digital oblivion may conflict with the "historical interest of keeping and archiving present information." Data that is put online now can be a rich source for future generations to study "those of us who live at the dawn of the digital age" [35].

Social Pressure on Industry and Government

Citizens can use accessible digital records to check whether their political representatives act appropriately. Politicians can utilize information found online to forecast general trends and developments. Those forecasts can help politicians to adjust policies and regulations before issues arise [2, 3].

Consumers can inform themselves online about businesses, services, and products. This information can be used to impose pressure upon companies and industries [2, 3]. Mayer-Schönberger [3] points out that "public shaming [of companies polluting the surrounding or not meeting certain standards, e. g., hygiene standards,] does have an impact on industry behavior."

Aiding Remembering and Learning From the Past

Mayer-Schönberger [3] states that "learning from history requires a societal capacity to remember." Digital memory counters human forgetfulness.

Individuals can use digital memory for everyday enhancements. Such enhancements can be reminders about special occasions like birthdays or anniversaries. Technology can be used to capture and revisit moments that are associated with joy and fulfillment. Family history can be preserved and friends one lost track of can be remembered [3, 45].

On a societal level, "[t]here might be a great public interest in the remembrance of information" [39]. Remembering can help society as a whole to learn from past mistakes. Comprehensive memory prevents that information becomes incomplete or misrepresentative of reality [3, 39].

Saving Lives and Preventing Harm

There are situations where large amounts of information about a person or an event can be helpful. Digital medical files and records of past illnesses, accessible for doctors and health personnel, can prove life-saving for a person. Recorded data about incidents and accidents are far more accurate than the reconstruction of a witness. This accuracy helps investigators and might lead to improvements in safety and security. The security of individuals can be increased by data logging and data preservation. Potential criminals might not carry out their illegal deeds when they know that their doings could be recorded [3, 45].

Innovation and Economic Growth

Large amounts of accumulated data lead to innovation and fuel economic growth [3]. Kwak et al. [31] state that "[b]y utilizing myriad data, enormous business opportunities have emerged, which has unleashed a huge wave of innovation." Products utilizing personal information of everyday situations can improve the quality of life for an individual or a group.

Market Analysis and Personalized Content

Utilizing large amounts of (personal) data enables businesses and governments to design, produce, and provide highly personalized goods and services. It allows a more precise marketing, sparing those who are not the target group of a good or service unwanted advertisement. This results in a better service for customers and citizens, improving their quality of life [3, 31, 68].

Ensuring Social Qualifications of Employees

Bishop et al. [2] point out that information found online can be used by companies to check the social qualifications of a job applicant. Recruiters can check whether applicants make negative statements about previous employers, check legal records, see if applicants engage in drunkenness, and if other professionals admire or criticize them.

2.2 Definition of Digital Oblivion

This section defines the term *digital oblivion* as it is understood in the following chapters of this thesis.

Definition 2.1 (Digital Oblivion (General))**.** *Digital oblivion summarizes the transfer of forgetting to the digital world. It is the state of digital data being forgotten, especially by the public.*

The following chapters focus on digital oblivion in the context of OSN.

Definition 2.2 (Digital Oblivion (OSN))**.** *In the context of online social networks (OSN), digital oblivion is the state of the following data about an individual being unavailable, removed from the OSN, or non-linkable to the individual:*

- *Data uploaded by the individual herself,*
- *Data uploaded by other users about the individual,*
- *Data linking the individual to still available information (e. g., to the authorship of a text),*
- *Contextual data which allows the reconstruction of the original data.*

This data includes images depicting the individual, texts where the individual is publisher, author or co-author of, tags which link the individual to content, posts of other users the individual shared, posts of other users about the individual, and any other content uploaded by the individual.

3 Method

This chapter introduces the study design. It starts with the introduction of the research question and continues with details on the survey. This chapter ends with information on the demography of participants and how participants were chosen.

This thesis seeks to answer the following question by conducting a user study:

> *Is the absence of forgetting mechanisms in online social networks (OSN) a problem for users?*

Additionally, this study is designed to evaluate against who or what users want to defend themselves by enabling digital oblivion in OSN. This question is not the main focus of this study. Hence it is not evaluated in detail or depth.

3.1 Survey Design

This section begins with a summary of the technical implementation of the online survey. The second part of this section describes how the survey is structured and contains information on the questions included in it.

The survey was implemented by using the survey framework LimeSurvey[1] (Version 3.17.7). It was hosted on a subdomain of Ruhr University Bochum (RUB), namely `www.mobsec-studies.rub.de`. The visual appearance of the survey was adapted to the RUB Corporate Design guidelines[2].

The survey is divided in 11 content groups. Each group has a specific purpose, e. g., informing the participant or asking certain types of questions. The full survey and all questions can be found in Appendix A.

1. Introduction, Data Protection and Consent The first section of the survey includes information about the intention of the study, which data is collected, that data is processed anonymously, and how the data will be used. The participants were asked to agree to these terms before proceeding. Texts in this section were formulated in a way that avoids the priming of participants.

[1] `https://www.limesurvey.org`
[2] `https://www.ruhr-uni-bochum.de/cd`

2. Demography – Part 1 This question group consists of general demographic questions like the participants' age, gender, and the country they live in.

3. General Questions on Social Media This question group starts with a short explanation of the terms "social media" and "social networks" (following the definitions found in [69–71]). Afterwards, the participants were asked questions on how and how frequently they use social media.

The intention of this question group is to have an indicator of how much own experience the participant has with OSN. Additionally, free text responses might indicate if inactivity in OSN is related to privacy issues or to mechanisms implementing digital oblivion being absent.

4. Scenarios Five scenarios were presented to the participants. They were asked to answer according to their own experiences of such a scenario. If no personal experience exists, participants were asked to imagine how they would react in this situation. The scenarios were displayed in random order to evade learning effects.

Scenario 1: Image Reference

> *Imagine you published an image on Facebook. Later you deleted the post including the image. After the post was deleted, you find copies of your image within the posts of other people. You do not know if someone actively distributes a screenshot of your image or if this is a reference on your deleted post still displaying the image.*

Participants were asked whether they consider this problematic and if they already made this experience in a social network.

The intention here is to indicate whether the participants are aware of this kind of situation and if they consider it problematic. Additionally, the responses in this scenario point out whether a participant has some first-hand experience with this kind of situation.

Scenario 2: Screenshot Sharing

> *You uploaded and shared an image on Instagram. The picture is neither embarrassing for you nor does it depict sensible information, yet you do not want this image to be shared publicly. You have taken appropriate actions to ensure this, for example you restricted the visibility or deleted the image some time later.*

> *Now you discover that someone has taken a screenshot of your post and shares this screenshot publicly in the social network.*

First, participants were asked whether they would actively try to restrict or prevent the distribution of the screenshot.

If the participants decided to restrict the distribution of the screenshot, they were asked how important certain types of support are for them, e. g., support by family, friends, official support of the OSN, or tools which help them to find, report, or delete all occurrences of the screenshot. In addition, participants were asked if their motivation to prevent the distribution of the screenshot depends on the person who shared the screenshot.

If the participants decided not to restrict the distribution of the screenshot, they were asked why they would not prevent it. The following questions aimed to find out if active support by certain groups of people or tools implementing digital oblivion would change their decision.

The intention behind this scenario is to find out whether the participants would wish for (automated) tools to aid the process of digital oblivion in this situation. The question about their motivation might be an indicator of users being particularly concerned about certain groups of people sharing information about them. In case the participant chose not to prevent or restrict the distribution of the screenshot, this question group aims to find out whether tools implementing digital oblivion would change the decision of the participant.

Scenario 3: Information Spreading Through Third Parties

> *You come across a public post on Facebook where the author is sharing information about you. You do not want this information to be publicly visible on Facebook; actually, this information should never be shared on any social network.*

Participants were asked the same questions as in the previous scenario (Screenshot Sharing). Additionally, all participants were asked whether they had experienced this situation in a social network. If they had experienced it, they were asked how they got to know about the post in which the information was shared.

The intention of this scenario is the same as for the previous scenario (Screenshot Sharing). This scenario presents a different situation as the participant did not initially share the information. The question of how participants got to know about the shared information was added to obscure the exact purpose of this study.

Scenario 4: Account Deletion

> *You published several tweets with your Twitter account and interacted with other members of the social network.*
>
> *Now you delete your Twitter account.*

Participants were shortly introduced to an automated tool that could verify if their tweets and conversations are not publicly available anymore. They were asked whether they would like to have such a tool.

Another question was if participants believe that their posts are still publicly available after they deleted their accounts. Also, participants were asked if they think that Twitter deletes their data from Twitter's servers. Depending on the answer to those two questions, participants were asked whether they consider this problematic.

This question group intends to ask whether the participants would like automated tools to verify that their content is removed from public view once they delete their accounts. Additionally, this question group aims to find out whether participants consider it problematic when some of their data is still publicly available or stored on the OSN provider's servers after they deleted their account. Another intention is to know how far participants trust the OSN provider regarding the deletion of data.

Scenario 5: Forgotten Image

> *On Instagram, one of your fans/followers asks you about an image. You vaguely recognize the image but you don't recall where you know it from. A short research shows: you uploaded this image on Instagram some time ago and forgot about this post over time.*

Participants were asked if they want their images to be displayed to them again after some time. As a second question, participants were asked if they want their posts to be automatically deleted after a while.

This question group covers the time dimension of digital oblivion and human forgetfulness. The first goal is to find out whether participants wish for automated support to be reminded of their content. The second goal is to evaluate if participants want their content to be automatically deleted.

5. Demography – Part 2 This question group asks questions regarding the profession and interests of the participants. This includes the highest level of education, the professional field participants are active in, and whether they engage privately or professionally in social media, IT Security, or data protection and data privacy.

The primary purpose of this question group is to know whether there is a bias in the group of participants.

6. End Message The survey ends with a *Thank you* message and the display of a contact email address.

3.2 Pilot Study

A pilot study was conducted before the survey was published for general participation. The goal of the pilot study was to find questions or answer options that lead to misunderstandings or misinterpretations, to find design weaknesses, and to improve the usability of the survey.

During the pilot study, the survey was given to one participant at a time. The participants filled out the survey in the presence of the survey designer and were asked to speak out thoughts or notes they had. Sometimes the survey designer asked questions, e. g., when the participant seemed unsure. Occasionally, participants were asked to choose one answer over another in order to evaluate a specific set of questions.

The pilot study was conducted within eight days, with six participants and had four distinct phases. After each phase, the feedback was implemented before the next phase started with other participants.

The following list summarizes the changes:

- Introduction of the *Back*-button.
- Implementation of a second progress bar right above the *Next*- and *Back*-button.
- Addition of the second Demography section.
- More detailed descriptions of several scenarios and questions.
- Changes in the question and answer order.
- Removing and adding of answer options.
- Splitting of the introduction section and one question in several parts.
- Changing and adjustment of Likert scales.
- Adjustment of wording and spelling.
- Introduction of distinct icons for each scenario.

3.3 Participants

First, this section describes how the participants were recruited. The second part of this section includes demographic information about the participants.

3.3.1 Recruitment

The survey was accessible for participation for three weeks. Participants were recruited through several channels:

Friends and Acquaintances Friends and acquaintances were asked to participate in the study. They were asked to spread the survey link wherever they think it is appropriate.

Mailing Lists The mailing lists of Ruhr University Bochum[3] were used to distribute the survey link.

Social Media A short information text, including the link to the survey, was distributed on several social media platforms.

To better spread the survey link and make it easier for others to share the link, a short information text was written in both German and English. For details on the information text, refer to Appendix B.

No monetary compensation was offered to the participants for participation in the study.

3.3.2 Exclusion of Participants

257 participants completed the survey. Of those, 4 participants were younger than 18 years. Due to privacy considerations, people who were younger than 18 years of age were excluded from the study.

The received datasets were evaluated for the time the participants needed to complete the survey. The mean time to complete the survey is 10 minutes. All datasets where the participant needed less than 6 minutes (34 participants) or more than 15 minutes (25 participants) were manually checked for completeness. Two datasets were removed during this step.

Datasets were investigated on whether a participant always chose the same answer options. This was checked to remove datasets where the participant only clicked through the survey. No such dataset was found.

[3] `https://lists.ruhr-uni-bochum.de/mailman/listinfo`

Datasets were looked through on whether participants skipped whole scenarios. If a participant skipped only one scenario and seemed to have answered the other scenarios seriously, the dataset was not removed. These datasets were kept because the answers for the other scenarios are still relevant. Two participants skipped one scenario. One dataset was removed because the participant had skipped two scenarios.

The data analysis in this and the following chapters includes 250 datasets.

Services that do not fit into the definition of social media given in the survey were removed from free text responses. An example of such a service is the messaging service WhatsApp[4].

Free text responses that are equivalent to given answer categories were deleted. It was made sure that the relating answer category was chosen instead.

3.3.3 Demography

This section presents the demographic traits of the participants, their education, and fields of profession and interest.

Age and Gender

Participants are between 18 and 73 years old. Almost 50 % are younger than 30 years, and 70 % are younger than 40 years. On average, participants are 35 years old; the median age is 30. The age distribution is displayed in Table 3.1.

Table 3.1: Age distribution of survey participants. On average, participants are 35 years of age.

Age	Participants	Percentage
18-19	6	2.4 %
20-24	68	27.2 %
25-29	50	20.0 %
30-35	40	16.0 %
36-39	11	4.4 %
40-49	20	8.0 %
50-59	29	11.6 %
60-73	26	10.4 %
Total	250	100 %

[4] `https://www.whatsapp.com`

The distribution between male and female participants is almost equal. As can be seen in Table 3.2, there are more male participants than female participants. Of those choosing neither male nor female as gender, 3 participants chose *other* and 4 participants decided not to disclose their gender.

Table 3.2: Gender distribution of survey participants. The distribution between male and female participants is almost equal.

Gender	Participants	Percentage
male	128	51.2 %
female	115	46.0 %
other, not disclosed	7	2.8 %
Total	250	100 %

Geographic Distribution

239 participants (95.6%) are living in Europe, 8 participants in the United States of America, 1 in Mexico, 1 in Australia, and 1 in Africa (who did not disclose the country).

As can be seen in Table 3.3, all except two participants living in Europe are from German-speaking countries (Germany, Austria) or countries where German is one of several official languages (Switzerland, Luxembourg).

Table 3.3: Residence distribution within Europe. Most participants are living in German-speaking countries.

Country	Participants	Percentage (Europe)	Percentage (all participants)
Austria	24	10.0 %	9.6 %
Germany	210	87.9%	84.0 %
Luxembourg	1	0.4 %	0.4 %
Netherlands	1	0.4 %	0.4 %
Switzerland	2	0.8 %	0.8 %
United Kingdom	1	0.4 %	0.4 %
Total	239	99.9 %[a]	95.6 %

[a]The deviation from 100 % can be explained by rounding errors

Education and Fields of Profession

Almost 91% of the participants received a school leaving qualification. More than half of the participants earned at least one degree at university and about 34% earned a Master's degree or equivalent (see Table 3.4).

More than half of the participants engage in IT, engineering, technique, or mathematics. About 10 % engage in medicine, health, and psychology and about 9% engage in education and social fields (see Table 3.5).

Table 3.4: Highest level of school or degree participants completed or received. More than half of the participants earned at least one degree at university.

Degree	Participants	Percentage
No school degree	0	0 %
Less than high school degree or equivalent	5	2.0 %
Finished vocational training	18	7.2 %
School leaving qualification	77	30.8 %
Bachelor degree	63	25.2 %
Master degree/graduate degree	70	28.0 %
Doctor degree	14	5.6 %
Other	3	1.2 %
Total	250	100 %

Table 3.5: Fields of profession of the participants' current activity.

Profession	Participants	Percentage
Administration, Management, Law	15	6.0 %
Art, Culture, Literature	10	4.0 %
Economic Sciences	4	1.6 %
Education, Social	23	9.2 %
IT, Engineering, Technique, Mathematics	131	52.4 %
Media, Communication, Advertisement	19	7.6 %
Medicine, Health, Psychology	26	10.4 %
Natural Sciences, Life Sciences	2	0.8 %
Other or not disclosed	20	8.0 %
Total	250	100 %

Study-Related Private or Professional Interests

About 16% of participants engage professionally in social media, while more than half of the participants do not engage at all with social media professionally (see Table 3.6). About half of the participants engage in social media privately whilst about 16% do not engage in social media privately (see Table 3.7). This indicates that the participants use social media far more often in private context than in their professional context. 29 participants (11.6 %) engage professionally as well as privately in social media.

40% of the participants engage professionally in data protection and data privacy. Together with those participants engaging a little bit in this field, this makes about 75% professional engagement (see Table 3.6). About 38% of the participants engage privately in data protection and data privacy, about 43% do so a little bit (see Table 3.7). 64 participants (25.6 %) engage professionally and privately in data protection and data privacy. This indicates a general interest and knowledge of the participants in these topics.

More than half of the participants engage professionally in IT security, another 24.1% engage a little bit in this field (see Table 3.6). About 41% of the participants engage privately in IT security, while about 39% do so a little bit (see Table 3.7). This indicates that about 70-80% of the participants have some knowledge and engagement in IT security. 78 participants (31.2 %) engage professionally and privately in IT security.

Table 3.6: Percentage of participants who engage professionally in different study-related fields.

	Yes	A little bit	No	Not answered
Social Media	15.8 %	31.6 %	52.2 %	0.4 %
Data Protection and Data Privacy	40.0 %	35.2 %	23.3 %	1.6 %
IT Security	50.6 %	24.1 %	23.7 %	1.6 %

Table 3.7: Percentage of participants who engage privately in different study-related fields.

	Yes	A little bit	No	Not answered
Social Media	47.4 %	32.0 %	16.2 %	4.4 %
Data Protection and Data Privacy	38.3 %	43.1 %	14.6 %	4 %
IT Security	40.7 %	39.1 %	16.2 %	4 %

4 Results

This chapter presents the results of the study. The first section describes the methods used for presentation and evaluation of the study. The following sections each present the results of one scenario or question group of the survey (for an overview on the survey, see section 3.1). Sections presenting the results of one scenario begin with the description text of the scenario. The results are discussed in chapter 5.

4.1 Statistics Used for Data Analysis and Presentation

The evaluation of this study utilizes quantitative methods. This means that the results are presented and analyzed numerically.

Results in this chapter are presented by the means of descriptive statistics. This allows for a clear, structured, and compact representation of the data [72]. This representation is then used to interpret the data in chapter 5.

The following representations are used in this chapter:

- Frequency tables are used to present data where more than three different response options were presented to the participant. Also, frequency tables are used to present free-text responses of participants.

- Histograms are mostly used to present responses to questions where up to four distinct response options were presented to the participant.

- Median, means, and quartiles are used and represented as box plot. In this thesis, box plots are used to visualize responses basing on a Likert scale. This allows for a visual representation of tendencies in the responses.

In addition, the findings are summarized as text. Textual representation of the data is also used to point out certain findings.

4.2 Inactivity and Non-Activity in Online Social Networks

At the beginning of the survey, participants were asked about their activity in online social networks (OSN). Participants who indicated that they are no active users of OSN were asked for their motives for not being active users of OSN. All other participants were asked if there are social media services that they do not use anymore and what caused their inactivity.

15 participants commented on why they do not use OSN. Four of them pointed out privacy concerns as a reason. This makes privacy concerns the third-most named reason for not using OSN. A list of all reasons that participants gave as response to this question can be found in Table C.1 in Appendix C.

84 participants commented on the question of why they stopped using some OSN. 19 participants pointed out privacy issues and worries about how the service provider handles their data. This makes it the third-most named reason. A list of all reasons that participants mentioned can be found in Table C.2 in Appendix C.

4.3 Scenario 1: Image Reference

> *Imagine you published an image on Facebook. Later you deleted the post including the image. After the post was deleted, you find copies of your image within the posts of other people. You do not know if someone actively distributes a screenshot of your image or if this is a reference on your deleted post still displaying the image.*

20 participants (8 %) stated that they had experienced this situation in an OSN.

171 participants (68.4 %) consider it problematic when their image is still visible or distributable once that image has been deleted. 57 participants (20.8 %) point out that it strongly depends on the content of the image. 21 participants (8.4 %) state that they do not have a problem when their images are still visible or distributable after they deleted them (see Table 4.1).

4.4 Scenario 2: Screenshot Sharing

> *You uploaded and shared an image on Instagram. The picture is neither embarrassing for you nor does it depict sensible information, yet you do not want this image to be shared publicly. You have taken appropriate actions to ensure this, for example you restricted the visibility or deleted the image some time later.*

Table 4.1: Responses to the question if participants consider it problematic when an image they deleted is still visible. The majority of participants (68.4 %) consider this problematic in any or in most cases.

	Participants	Percentage
Yes, in any case	119	47.6 %
Yes, in most cases	52	20.8 %
It strongly depends on the content	57	22.8 %
No, I usually have no problem with that	11	4.4 %
No, I never have a problem with that	10	4.0 %
Not answered	1	0.4 %
	250	100 %

Now you discover that someone has taken a screenshot of your post and shares this screenshot publicly in the social network.

Participants were asked whether they try to prevent or restrict the distribution of the screenshot. 174 participants (69.6 %) decided to prevent or restrict the distribution. 76 participants (30.4 %) chose not to prevent or restrict the distribution of the screenshot. One of the participants who decided to prevent the distribution skipped the remainder of the questions in this scenario.

4.4.1 Participants Trying to Prevent or Restrict the Distribution

Participants were asked how important active support by different groups of people or tools is for them in this situation (see Figure 4.1). Being actively supported by the official support of the OSN is considered *very important* by 77.5 % of the participants and *important* by 16.2 % of the participants. This makes the official support of the OSN the most important active supporter, followed by the three tools implementing aspects of digital oblivion. Of those three tools, the one finding and showing each occurrence of the screenshot to the participant is the most important one. Active support by friends is considered as *important* by the participants, but in comparison with support by the three tools implementing digital oblivion, participants chose the options *of little importance* and *not important at all* more often. Participants consider support by strangers or by their followers/fans as least important, or do not wish to be supported by these groups of people at all.

14 participants responded to the question whether there are other people, groups, or tools they would like to be actively supported by in this situation. Most often, the response to this question was public authorities or lawyers and legal helplines. Both were mentioned by four participants. Three participants would like active support by the person who shared the screenshot or the social environment of that

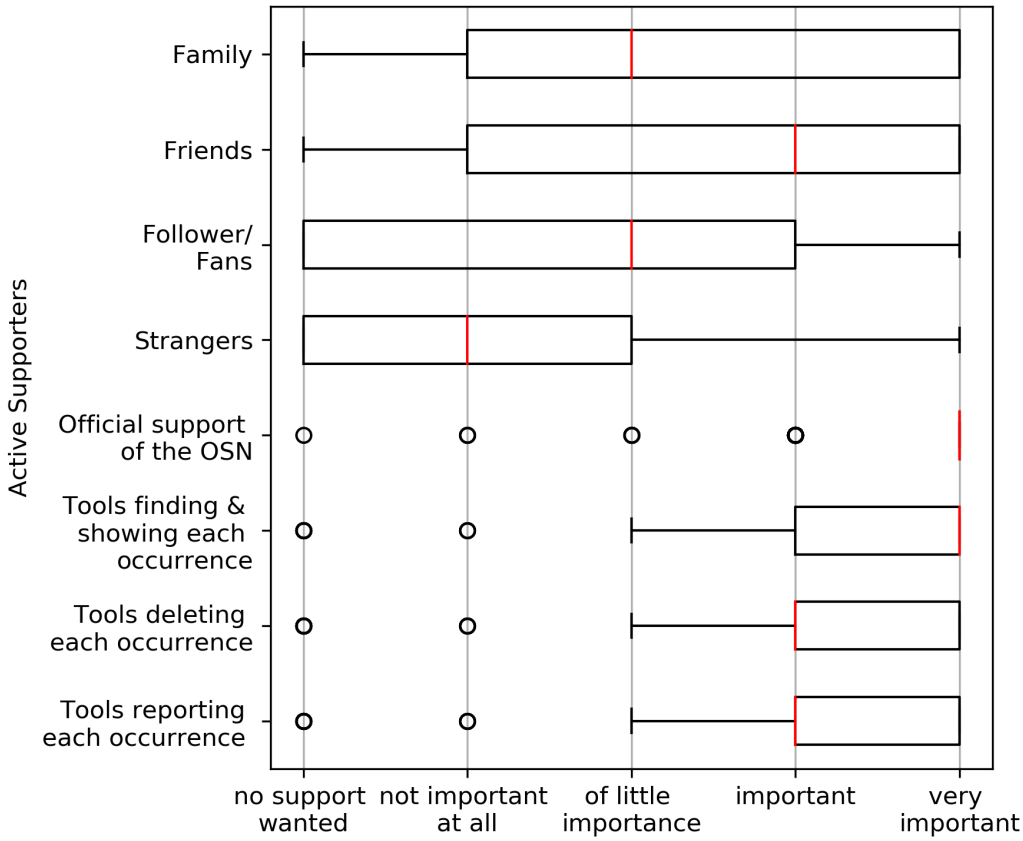


Figure 4.1: Importance of active support by certain groups of people and tools to prevent or restrict the distribution of the screenshot. The red bar denotes the median importance. Participants consider support by the official support of the OSN as most important, followed by the three tools implementing aspects of digital oblivion.

person. Another two participants responded that they wish for assistance from the police.

The motivation to prevent or restrict the distribution of the screenshot is between medium and high for all groups of people sharing the screenshot (see Figure 4.2). It can be noted that there is a trend towards medium motivation for family members and friends. For other acquaintances, followers/fans, fellow workers, superiors, and strangers, there is a trend towards high motivation. The median motivation to prevent or restrict the distribution of the screenshot is high for all groups of distributors.

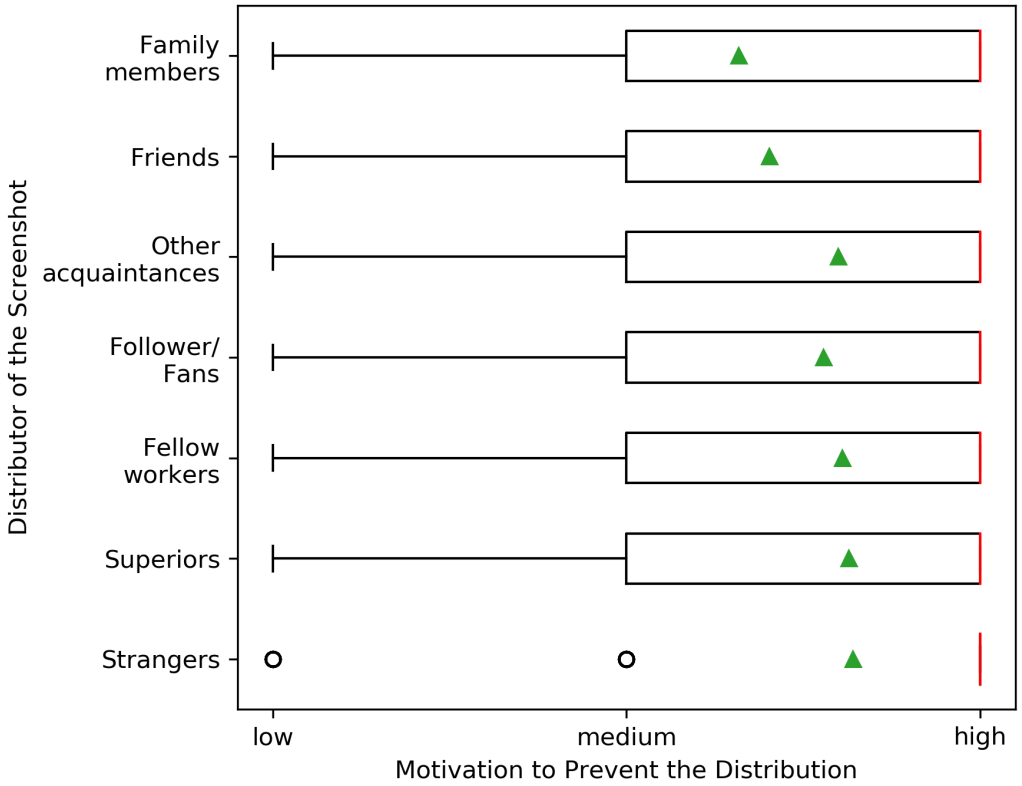


Figure 4.2: Motivation to prevent the further distribution of a screenshot that had been shared by others. The red bar presents the median motivation; the green triangle presents the mean motivation. The motivation is between medium and high for all groups of people, leaning towards a high motivation in most cases.

4.4.2 Participants Not Trying to Prevent or Restrict the Distribution

The participants pointed out several reasons why they do not prevent or restrict the distribution of the screenshot (see Table 4.2). The most named reasons are that it is not important enough for the participants (53.9 %) or that they do not want to spend time on it (39.5 %). A closer inspection of data shows that participants choosing one of these reasons also chose other reasons why they do not prevent the distribution of the screenshot. It can be assumed that the reason depends on what is depicted by the screenshot. Other reasons that were chosen by more than 22 % of the participants were not knowing how to find all occurrences of the screenshot, that an action on their side would not change anything, or that it is actually good that the screenshot was shared.

Table 4.2: Reasons why the participants do not prevent or restrict the distribution of the screenshot. Multiple selections were possible. The two most named reasons are that this is not important enough for the participants or that they do not want to spend time on it.

	Participants	Percentage
Not important enough	41	53.9 %
Do not want to spend time on it	30	39.5 %
Do not know how to contact the person sharing the screenshot	6	7.9 %
Support team of the OSN is not helpful	6	7.9 %
Do not know how to find all occurrences of the screenshot	22	28.9 %
It would not change anything	20	26.3 %
It would only make things worse	1	1.3 %
It is good that the screenshot was shared	17	22.4 %
Other	6	7.9 %

Participants were asked if active support by certain groups of people or tools would change their decision (see Figure 4.3). Only active support by the official support of the OSN would motivate more than half of the participants to prevent or restrict the distribution of the screenshot. Active support by the three tools implementing facets of digital oblivion would change the decision of 30-48 % of the participants, depending on the tool. Active support by different groups of people has the potential to change the decision of less than 30 % of the participants. Of those groups, active support by friends has the most potential; it would motivate 26.3 % of the participants to actively prevent or restrict the distribution of the screenshot.

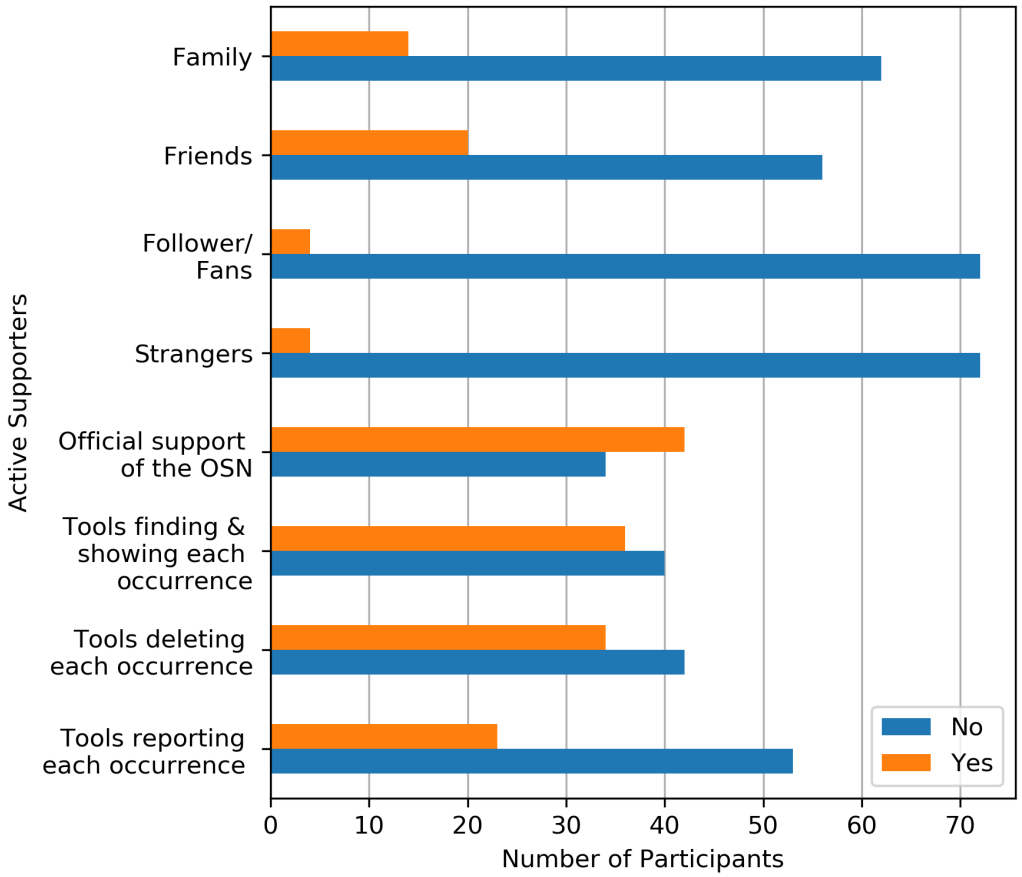


Figure 4.3: Responses whether active support by certain groups of people or tools would change the participants' decision, resulting in them actively trying to prevent or restrict the distribution of the screenshot. Only active support by the official support of the OSN would change the decision of more than half of the participants.

4.5 Scenario 3: Information Spreading Through Third Parties

> *You come across a public post on Facebook where the author is sharing information about you. You do not want this information to be publicly visible on Facebook; actually, this information should never be shared on any social network.*

Participants were asked whether they try to prevent or restrict the distribution of the information. 224 participants (89.6 %) decided to so. 26 participants (10.4 %) chose not to prevent or restrict the distribution of the information. One of the participants who decided to prevent the distribution skipped the remainder of the questions in this scenario.

4.5.1 Participants Trying to Prevent or Restrict the Distribution

Participants were asked how important active support by different groups of people or tools are for them in such a situation (see Figure 4.4). Support by the official support of the OSN is most important, followed by two types of tools implementing digital oblivion:

1. tools that automatically report each occurrence of the information to the official support of the OSN and
2. tools that automatically find each occurrence of the information and show it to the user.

Active support by friends is considered as *important* by the participants. In comparison with support by the two tools implementing digital oblivion, support by friends was considered *of little importance* and *not important at all* more often. Participants consider support by their followers/fans and by strangers as least important or do not wish to be supported by these groups of people at all.

15 participants responded to the question if there are other people, groups, or tools they would like to be actively supported by in this situation. Most often, participants responded to this question with public authorities (5 participants), followed by administrators or moderators of the OSN (4 participants). Three participants would like to be actively supported by the person who shared the information or the social environment of that person. Another three participants would like to be supported by their lawyers or by the public prosecution department. Two participants responded that they wish for assistance from the police.

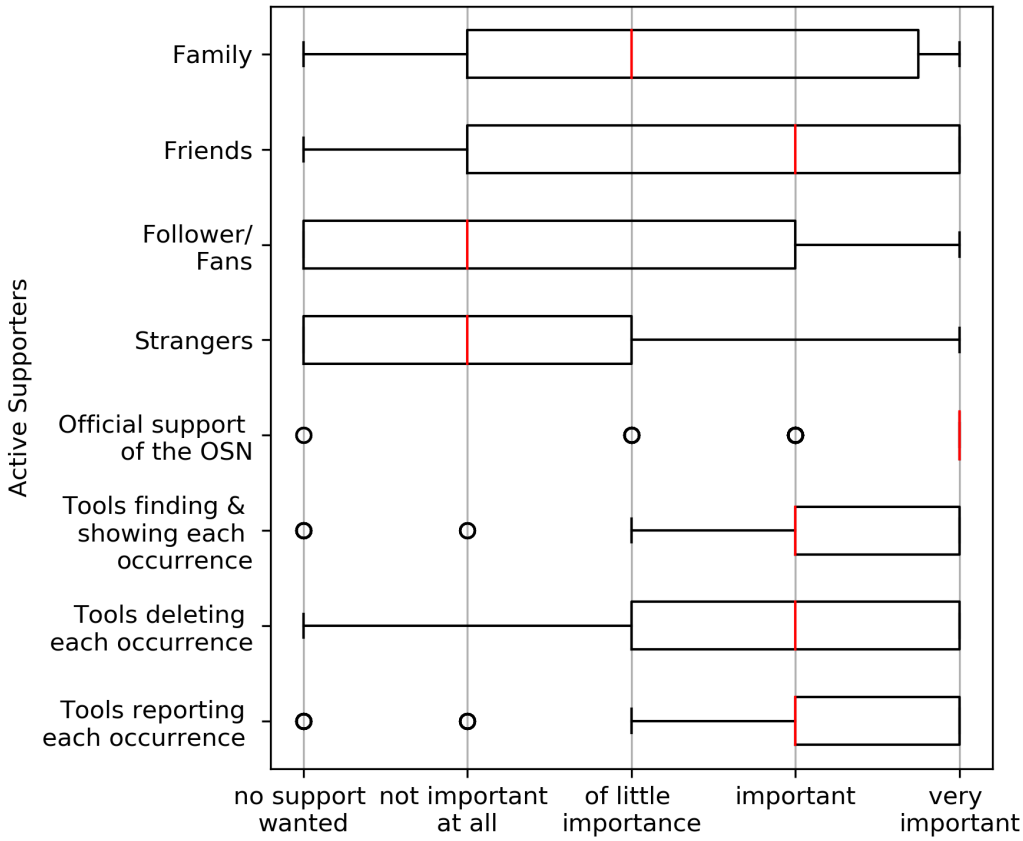


Figure 4.4: Importance of active support by certain groups of people or tools to prevent or restrict the distribution of information that was shared by others. The red bar denotes the median importance. Participants consider support by the official support of the OSN as most important, followed by two of the tools implementing aspects of digital oblivion.

The motivation to prevent the distribution of the information is between medium and high for all people who might have shared the information (see Figure 4.5). It can be noted that the trend is towards high motivation for all distributors except for family members. The median motivation to prevent or restrict the distribution of the information is high for all groups of people.

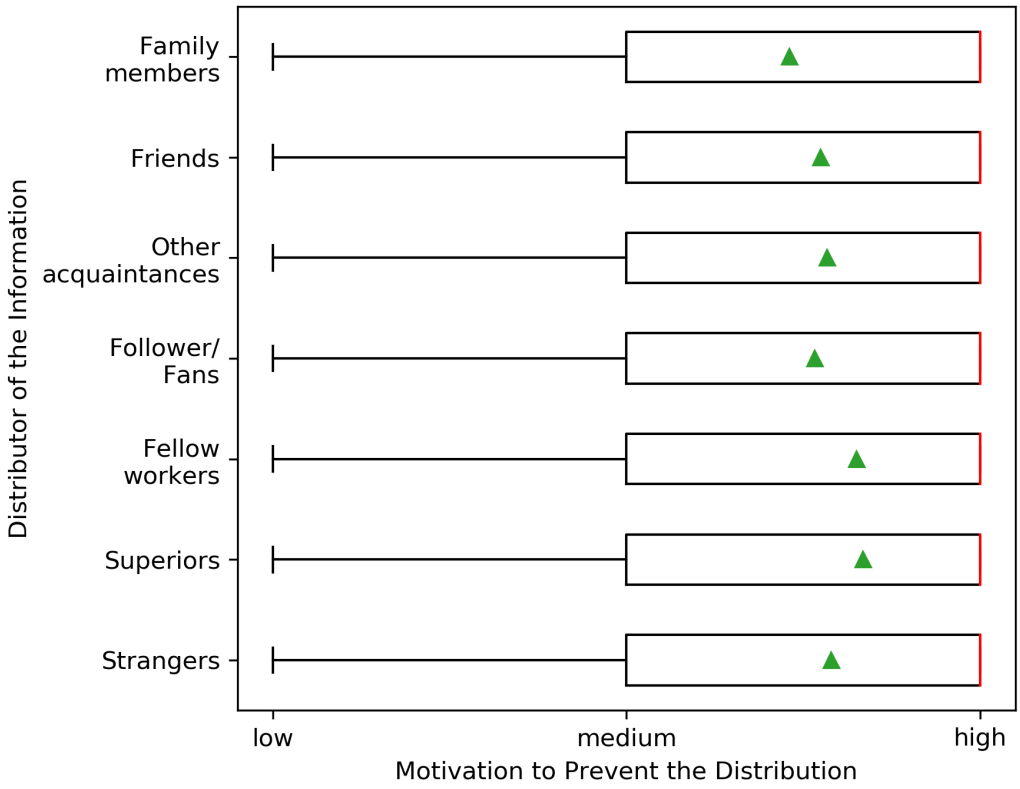


Figure 4.5: Motivation to prevent further distribution of information that was shared by others. The red bar presents the median motivation; the green triangle presents the mean motivation. The motivation is between medium and high, leaning towards a high motivation in all cases but family members.

4.5.2 Participants Not Trying to Prevent or Restrict the Distribution

The participants chose different reasons why they do not try to prevent or restrict the distribution of the information (see Table 4.3). Almost half of the participants state that it is not important enough for them. A closer inspection of the data shows that participants who chose this reason also chose other reasons why they do not prevent the distribution of the information. The next most important reasons are that an action by the participant would not change anything, that they do not want to spend time on this, and that it was actually good that the information was shared. No participant believed that actively preventing the distribution of the information would make the situation worse.

Table 4.3: Reasons why the participants do not prevent or restrict the distribution of the information. Multiple selections were possible. The two most named reasons are that this is not important enough for the participants or that an action on their side would not change anything.

	Participants	Percentage
Not important enough	12	46.2%
Do not want to spend time on it	9	34.6%
Do not know how to contact the person sharing the information	1	3.8%
Support team of the OSN not helpful	1	3.8%
Do not know how to find all occurrences of the information	4	15.4%
It would not change anything	10	38.5%
It would only make things worse	0	0%
It is good that the information was shared	6	23.1%
Other	3	11.5%

Participants were asked if active support by certain groups of people or tools would change their decision (see Figure 4.6). Active support by the official support of the OSN would result in 57.7 % of the participants trying to prevent or restrict the distribution of the information. A tool that automatically reports each occurrence of the information to the official support of the OSN would change the decision of half of the participants.

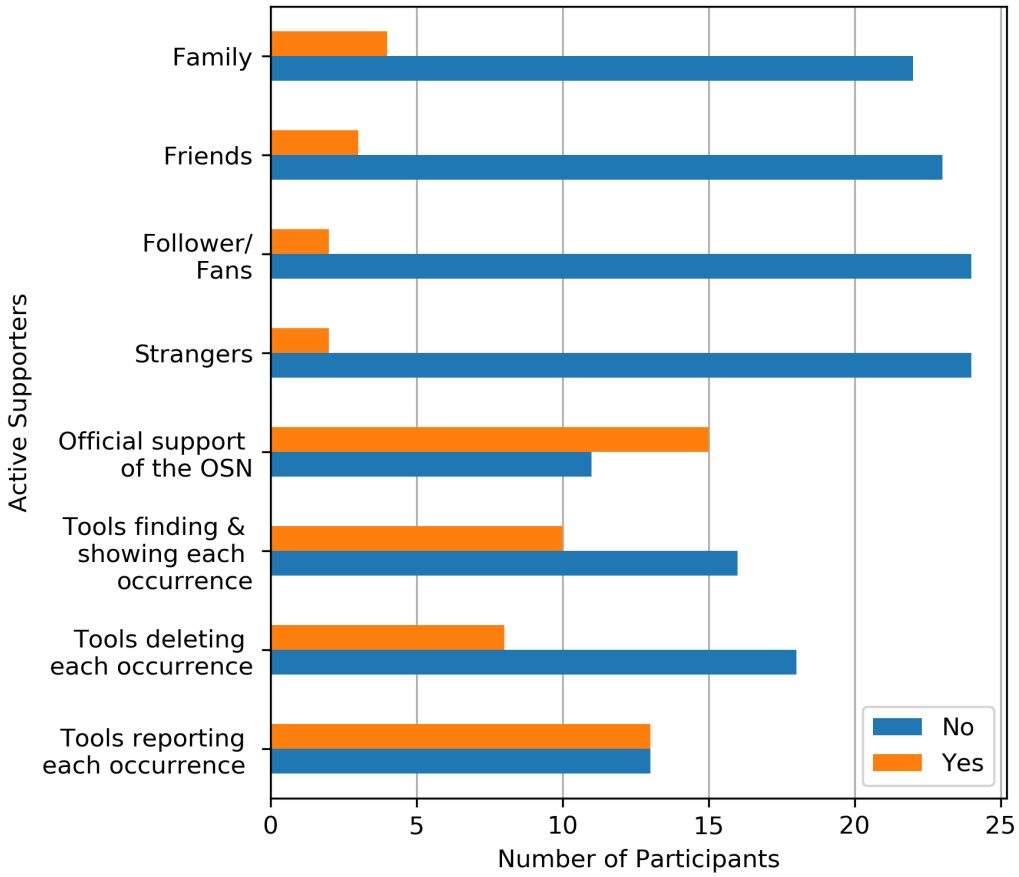


Figure 4.6: Responses whether active support by certain groups of people and tools would change the participants' decision, resulting in the participants' actively trying to prevent or restrict the distribution of the information. Active support by the official support of the OSN would change the decision of 15 participants. A tool that automatically reports each occurrence of the information to the official support of the OSN would change the decision of 13 participants (50 %).

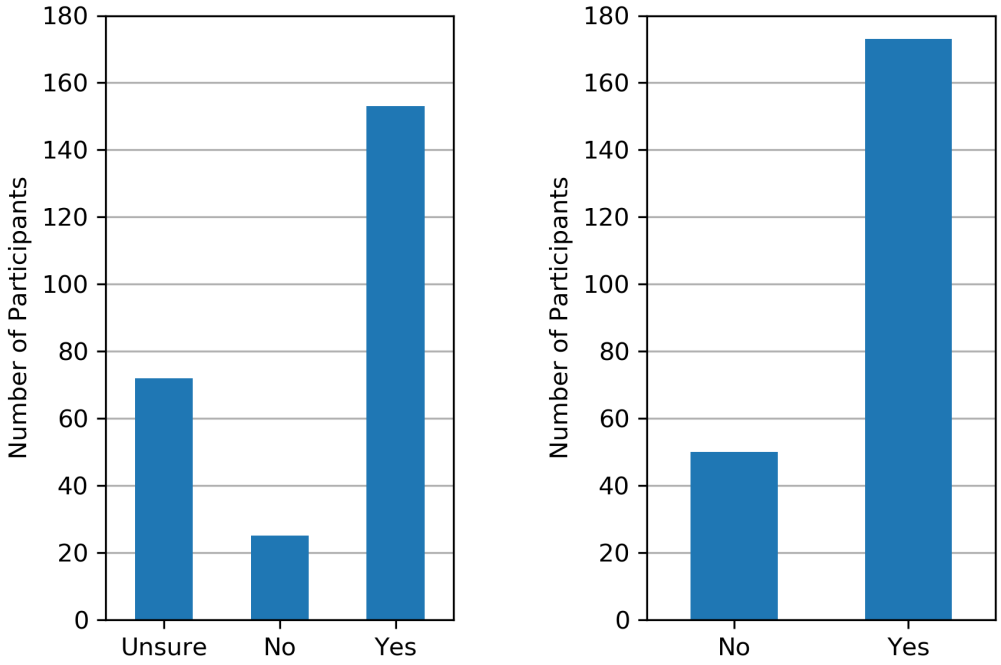
4.6 Scenario 4: Account Deletion

> *You published several tweets with your Twitter account and interacted with other members of the social network.*
>
> *Now you delete your Twitter account.*

Participants were asked whether they believe that some of their posts and conversations are still publicly available after they deleted their account. 90 % of participants are either unsure or believe that some of their posts are still publicly available (see Figure 4.7(a)). Those participants were asked if they would consider it problematic if their content was still publicly available. 173 participants (69.2 %) consider this problematic, 50 participants (20 %) responded that they do not consider it problematic and 2 participants chose not to answer this question (see Figure 4.7(b)).

206 participants (82.4 %) stated that they want a tool that verifies that all their posts and conversations are not publicly available anymore after they deleted their account. 42 participants (16.8 %) answered that they do not want such a tool, and 2 participants chose not to answer this question.

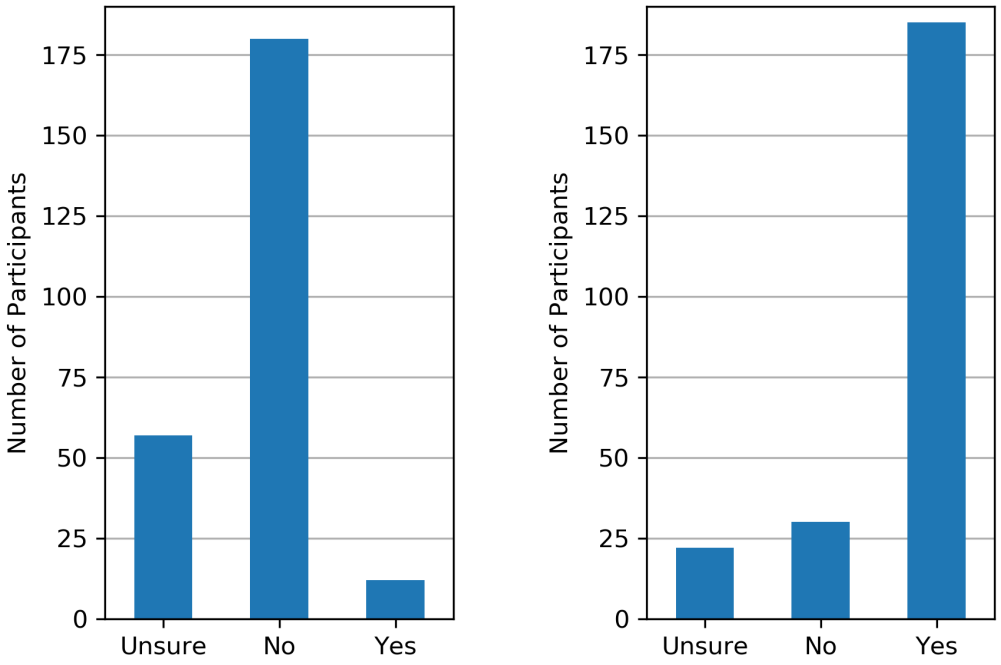
Participants were asked whether they believe that the service provider of the OSN deletes their content from the service provider's servers shortly after they deleted their account. 94.8 % of the participants believe that the service provider keeps their data or are unsure about it. 12 participants believe that the service provider does not keep their content on their servers after the participants deleted their account, and 1 participant chose not to answer this question (see Figure 4.8(a)). Those who believed that the service provider keeps their data or were unsure about it were asked whether they consider this problematic. 185 participants (74 %) stated that they consider it problematic, and 22 participants (8.8 %) were unsure. 30 participants (12 %) responded that they do not consider it problematic when their content stays long-term on the service provider's servers (see Figure 4.8(b)).



(a) Participants' believe whether content is still publicly available after they deleted their account.

(b) Participants consider it problematic if their content was still publicly available after they deleted their account.

Figure 4.7: Figure (a) displays whether participants believe that their content is still publicly available after they deleted their account. 90 % of participants believe so or are unsure. Figure (b) displays the responses to whether participants consider it problematic should their content be publicly available after they deleted their account. 173 participants (69.2 %) consider this to be problematic.



(a) Participants' believe whether the service provider deletes their content shortly after they deleted their account.

(b) Participants consider it problematic if their content was stored long-term on the service provider's servers.

Figure 4.8: Figure (a) displays whether participants believe that the service provider deletes their content from their servers shortly after they deleted their account. Almost 95 % of the participants do not think so or are unsure. Figure (b) displays the responses to whether participants consider it problematic should their content be stored long-term on the service provider's servers. 82.8 % of participants consider this as problematic or are unsure.

4.7 Scenario 5: Forgotten Image

> *On Instagram, one of your fans/followers asks you about an image. You vaguely recognize the image but you don't recall where you know it from. A short research shows: you uploaded this image on Instagram some time ago and forgot about this post over time.*

26 participants (10.4 %) had experienced this situation themselves.

Participants were asked whether they want a setting in their accounts so that the images they had uploaded in the past will be shown to them again after a while. 80 participants (32 %) would want to have such an option. Of those, 71 participants (88.8 %) want to view their images again after a while to decide to either keep or delete them (see Figure 4.9).

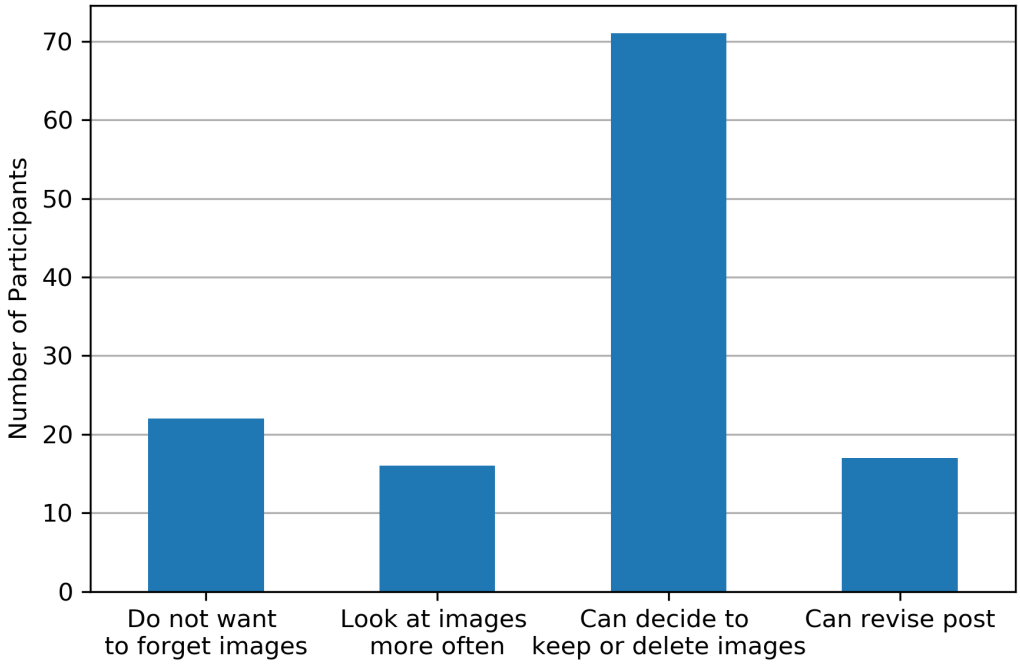


Figure 4.9: Responses to the question of why participants want to view images again that they posted a while ago. 80 of 250 participants chose that they would like to view old posts again after a while. 71 participants want to view their posts again after a while to decide to delete or keep them.

75.6 % of the participants do not want their content to be automatically deleted in OSN. 155 participants (62 %) stated that such a setting for automated deletion sounds good as an extra option, but they do not want it as a mandatory setting. 61 participants (24.4 %) want their content to be deleted automatically. Of those, 45

participants want to set the time until their content is deleted by themselves. 16 participants (6.4 %) want that the time until their content is deleted is set by the provider of the OSN (see Table 4.4).

Table 4.4: Participants were asked whether they want their posts to be deleted automatically after a fixed time. The majority of participants do not want the automated deletion of their posts.

	Participants	Percentage
Yes	16	6.4 %
Yes, and I want to set the time until the posts are deleted myself	45	18.0 %
No, but this sounds good as an extra option for each post	155	62.0 %
No	34	13.6 %
Total	250	100 %

5 Discussion

This chapter discusses the results of the study (for details on the results, see chapter 4).

The goal of this discussion is to answer the following research question: is the absence of forgetting mechanisms in online social networks (OSN) a problem for users? This question is answered in two ways. First, the data is analyzed on whether the presence of such mechanisms is wanted by users. Also, the data is analyzed on whether the presence of such mechanisms would enhance the users' experience in OSN. This is done because, as of today, such mechanisms are usually not implemented. Second, it is checked whether responses to the survey indicate that the absence of mechanisms implementing digital oblivion keeps users from engaging in OSN the way they would like to.

The first three sections of this chapter evaluate the research question based on different aspects of digital oblivion. The fourth section discusses trust in OSN providers and whether the motivation to prevent data spreading by others depends on the person sharing the data. The last section presents the limitations of this study.

5.1 Data Disclosed by Others

This section discusses the two scenarios when someone is sharing either a screenshot of a user's content or information about a user, both without the user's consent (see section 3.1 for details on the scenarios).

In this section, *data* refers to the shared screenshot as well as the distributed information of the two survey scenarios. Additionally, the term *data* is used in the common meaning of the word itself.

5.1.1 Awareness for Content Published by Users Themselves

Users are aware that data published by themselves online might be used in unintended ways. Accordingly, they react differently depending on what another person is sharing – content the users initially published themselves or information about the user that was not previously disclosed.

More than 69 % of the participants try to prevent or restrict the distribution of data that is shared by others without their consent. 20 % of the participants decided to prevent the distribution of information shared about them, but not the distribution of a screenshot depicting content that the participant shared initially.

Participants were asked why they decided not to prevent the distribution of the screenshot. Five of six free-text responses state that one should be aware that others might use content published online. Thus it would be their own fault when someone shares a screenshot of their content.

5.1.2 Support by the Official Support of the Online Social Network Is Most Important

Active support by the official support of the OSN is considered most important when someone else shares data belonging to or being about a user.

The official support of the OSN has the median importance *very important*. The majority of users classified support by the official support as *very important*. Additionally, no other supporter that was specified in the survey was marked as often as *very important* as the official support of the OSN.

5.1.3 Tools Implementing Digital Oblivion Would Be Appreciated

Users would appreciate the presence of tools implementing mechanisms of digital oblivion. Such tools may help them to handle data that is shared by others in an OSN in a better and more efficient way.

Participants consider tools implementing digital oblivion as *important* or *very important* when preventing the distribution of data shared by others. Only active support by the official support of the OSN is considered to be more important.

These findings are in line with the findings of Novotny et al. [10]. They found that users wish for oblivion due to privacy concerns over disclosed data and to have control over their data.

5.1.4 Users Want to Have Active Control Over Their Situation

Users prefer to use tools that allow them to have active control over their situation. They want to decide by themselves which steps to take next.

Three tools implementing aspects of digital oblivion were presented to the participants:

1. a tool that automatically finds each occurrence of the data and shows it to the participant,
2. a tool that automatically deletes each occurrence of the data, and
3. a tool that automatically reports each occurrence of the data to the official support of the OSN.

Of these three tools, participants consider a tool finding and showing each occurrence of the data to them as most important. A tool that automatically deletes all occurrences of the data is the least important one of the three tools, though it is also considered as important.

5.1.5 Importance of Active Support by People

Users consider active support by different groups of people as less important than active support by digital tools implementing digital oblivion.

Support by family members, followers/fans, and strangers has a median importance of *not important at all* or *of little importance*. In comparison, digital tools implementing digital oblivion have a median importance of *important* or *very important*.

Active support by friends has a median importance of *important*. This is the same median as for some of the tools implementing digital oblivion. The interquartile range of friends as supporters spreads over four categories, from *not important at all* to *very important*. In comparison, the interquartile range of supportive tools spreads over two categories – from *important* to *very important* – in all cases but one.

The participants were not asked why they consider one type of support to be more important than another. One can hypothesize that participants consider it harmful when certain people know about the data. Asking others for help might distribute the data in a way that participants wish to prevent. In comparison, the official support of the OSN has a well-recognized functionality regarding complaints and copyright issues. Tools or apps fulfill a predictable functionality and do not share data in unforeseeable ways. Thus, users might trust automated tools more in this situation than they trust people they know or do not know.

5.1.6 Support by Followers/Fans and Strangers Is Least Important

Users consider support by their followers/fans and strangers as least important or do not want to be supported by these groups of people at all.

The median importance of active support by followers/fans is *not important at all* or *of little importance*. The median importance of support by strangers is *not important*

at all in both scenarios. Both groups are the only supporters where the lower quartile is at *no support wanted*.

5.1.7 Digital Tools Have the Potential to Induce That Users Take Action

The presence of tools implementing digital oblivion in OSN has the potential to induce users to take active steps to prevent the distribution of data they do not wish to be shared or known.

About half of the participants would change their decision and actively try to prevent the distribution of data if they were supported by a tool implementing digital oblivion. Additionally, 15.4 %-28.9 % of the participants mentioned that they do not prevent the distribution in the first place because they do not know how to find all occurrences of the data. Here, automated tools could help users to find all occurrences of the data.

5.1.8 The Potential of a Tool to Change the Users' Decisions Depends on the Situation

Which of the tools implementing digital oblivion has the most potential to induce users to take action depends on the type of data that is shared against the users' will in the OSN.

When someone shares a screenshot of data that the user had shared initially, a tool finding each occurrence of the screenshot and showing it to the user has the highest potential to change the participants' decision. In case someone distributes information about the user, a tool reporting each occurrence of the information to the official support of the OSN has the highest potential to change the participants' decision.

Participants were not asked why they consider one tool to be more helpful than another. One can hypothesize that users contact the support team of the OSN more quickly when information is distributed about them than when their own content is distributed by someone else. Also, the support team has possibilities to penalize the behavior of the person sharing the information. In comparison, the sharing of content the user published initially might be seen as something one has to reckon with when one shares something in an OSN (also see subsection 5.1.1). Thus, it might be more important to know where the screenshot has been shared and to decide on the next steps based on this information.

5.1.9 The Official Support of the Online Social Network Has the Highest Potential to Induce That Users Take Action

Active assistance by the official support of the OSN has the highest potential to induce users to take active steps against the sharing of data without their consent.

Active support by the official support of the OSN would change the decision of more than half of the participants who decided to not prevent the distribution of the data in the first place. It is the only supporter that leads to more than half of the participants to take active steps against data being shared when they previously decided that they do not want to prevent the distribution.

5.2 Forgotten Content in Online Social Networks

This section discusses the case when users forget content that they posted in an OSN in the past.

5.2.1 Forgetting of Content Is Common

It is common for users in OSN to have forgotten some of the content they published online.

About 10 % of the participants state that they had forgotten an image they uploaded to an OSN and were reminded of this image by others. One can assume that some users forgot about images they uploaded to an OSN and were not reminded about it by others.

5.2.2 A Minority of Users Wants to Be Shown Their Old Posts Again

A minority of users would like to have a functionality or tool that shows their posts to them again after a while.

32 % of the participants indicated that they want their old posts to be automatically displayed to them again after a while.

5.2.3 A Tool Displaying Old Posts Would Be Used to Decide on Keeping or Deleting Data

A tool displaying old posts to a user would be primarily used to decide to either keep or delete the post. This means that such a tool is used to deliberately clean out the user's profile from time to time. Thus, such a tool would support the active enforcement of one aspect of digital oblivion in OSN.

Almost 90 % of the participants who want to have such a tool want to view their old posts again to decide to either keep or delete the post. Other selectable applications of such a tool – not wanting to forget their posts, wanting to look at their posts more often, and revising the post – were selected by 20-28 % of the participants.

5.3 Deletion of Content

This section discusses two topics related to deletion. The first three parts cover the automated deletion of content in OSN. The last two parts cover the deletion of an account.

5.3.1 Content Should Not Be Automatically Deleted

The majority of users does not want their content in OSN to be automatically deleted after a while.

75.6 % of the participants stated that they do not want that their content in OSN is automatically deleted after a while. This indicates that a general implementation of automated deletion mechanisms is not in the interest of the users. These findings are in line with the findings of Bauer et al. [11], Murillo et al. [12], and Ayalon et al. [15].

5.3.2 Optional Automated Deletion Would Be Appreciated

A majority of users would like to have an optional setting to delete their content after a fixed time automatically. Users may use this setting for some of their posts but not for all.

62 % of the participants consider an optional setting to delete their content after a fixed time a good idea. 24.4 % of the participants state that they want such an automated deletion as a mandatory setting for each of their posts. Together, these two groups make up 86.4 % of the participants.

5.3.3 Deleted Content Should Be Unavailable

Users want that content they deleted in an OSN is not available anymore.

68.4 % of the participants state that they consider it problematic when an image is still visible after they deleted it. 20.8 % of the participants pointed out that it depends on the content whether they consider this problematic. Almost 70 % of the participants consider it problematic when their posts are publicly available after they deleted their account.

5.3.4 Users Are Unsure Whether Content Is Unavailable After Account Deletion

Users are unsure whether their content is inaccessible by the public once they deleted their account. They consider this problematic.

90 % of the participants stated that they are unsure or believe that their content is still publicly available after they deleted their OSN account. Almost 70 % of the participants consider it problematic when their posts are available after they deleted their account.

5.3.5 Tools Checking for Data Availability After Account Deletion Would Be Appreciated

There is a need for tools that verify that no data is publicly available anymore in an OSN after users deleted their accounts.

Most participants are unsure or believe that their content might still be publicly available once they deleted their account. A majority of participants consider this problematic. This raises the need to check if their content is still publicly available after they deleted their account. 82.4 % of the participants stated that they would like to have an automated tool that verifies that all their posts and conversations are not publicly available anymore.

5.4 Adversaries and Trust in Social Network Providers

This section discusses two topics. The first part of this section discusses the question whether a user's motivation to prevent the distribution of data shared by others depends on the person sharing the data. The second part of this section discusses implications on whether users trust OSN service providers.

5.4.1 The Motivation to Prevent the Distribution of Data Is Independent of the Distributor

Someone shares data belonging to or being about a user in an OSN without the user's consent. The motivation to prevent the distribution of this data does not depend on the person sharing it.

The interquartile range of motivation to restrict or prevent the distribution of data belonging to or being about a user is between *medium* and *high* for all distributors. The median motivation is *high* for all distributors. While there are slight differences in motivation between different groups of people, one can not say that one group of people triggers an outstandingly high or low motivation to prevent the distribution of the data.

5.4.2 Users Believe That Service Providers Keep Their Data

Users believe that the OSN service provider keeps at least some of their content after the users deleted their accounts.

Almost 95 % of the participants believe that the OSN service provider keeps their posts and conversation on their servers or are unsure about it. These findings are in line with the findings of Murillo et al. [12], who found that users believe that some deleted data stays on the service provider's servers.

5.4.3 Trust in Service Providers

Users do not trust the OSN service provider to delete their data and handle their data properly once they deleted their accounts.

As pointed out in subsection 5.4.2, users believe that the service provider keeps at least some of the user's content on their servers. 74 % of the participants consider this problematic and another 8.8 % are not sure whether they consider this problematic. These findings emphasize the findings of Madden et al. [20], who found that the majority of users in the United States of America do not trust social network service providers.

There might be various reasons why the participants consider it problematic when the service provider keeps their data after they deleted their account. Mayer-Schönberger [3], Ambrose et al. [39], and Bode et al. [40] state that users do not know which data is kept by the service provider and that they are not in control of their data. Bishop et al. [2], Mayer-Schönberger [3], and Novotny et al. [10] point out that data that is kept by service providers might be used in unforeseeable ways. These insecurities can result in users not trusting the OSN provider. Additionally, one can hypothesize

that some users do not want that there exists a trace of them or their data linked to an OSN they are leaving behind.

5.5 Limitations

This section describes the limitations and biases of this study. It starts with the biases introduced by the participant sample and continues with design limitations.

5.5.1 Participants

This section describes limitations and biases introduced by the participant sample.

Age

All participants are 18 years of age or older. Thus, this study does not include viewpoints of teenagers or children.

As can be seen in Table 3.1, there is a bias in the age distribution. About half of the participants are between 18 and 29 years old.

Geographic Distribution

Most participants (95.6%) are living in German-speaking countries, 84 % are living in Germany. This indicates a strong geographical and cultural bias in the results of this study.

Bias in Education and Fields of Profession

As can be seen in Table 3.4, more than 90 % of the participants have finished high school or equivalent. Many of the participants having a school leaving qualification (30.8 %) might be enrolled at university. This assumption is based on the participants' age distribution and the knowledge that university channels or channels close to university were used to distribute the survey link. This indicates that the results of this study are biased in regards to the level of education and may have a tendency to represent views of academic professionals.

As can be seen in Table 3.5, 52.4 % of the participants engage in fields related to IT, engineering, technique, and mathematics. This presents a bias in regards to professions.

Bias Regarding IT Security and Privacy

As described in section 3.3.3, a high percentage of the participants engage professionally or privately in IT security, data protection, and privacy. It is not clear how in-depth the knowledge in this field is, yet the chances are high that this engagement introduces a knowledge or awareness bias into the study.

5.5.2 Self-Evaluation

When taking the survey, participants were confronted with certain situations and asked to self-evaluate how they would react and act in this situation. An observation of how participants would act in a real scenario within an OSN is not part of this study. There might be a difference in how participants think or say they will react and how they react when confronted with such a situation in the wild. In privacy research, this is known as "Privacy Paradox" [73–75].

5.5.3 Online Social Network Examples in Scenario Descriptions

Each scenario starts with the introduction of a specific OSN (Twitter[1], Facebook[2], or Instagram[3]). This specific introduction was chosen to make it easier for participants to imagine these scenarios and to keep the level of abstraction low.

The introduction of a specific OSN at the beginning of each scenario means that the results might be biased in regards to that specific OSN. For example, participants might trust Twitter to delete data from their servers when they delete their accounts. However, participants might not think the same about Facebook.

5.5.4 Asking for Absence and Not Presence of Mechanisms for Digital Oblivion

By design, this study evaluates if users see a problem in the absence of mechanisms for digital oblivion. Participants are asked whether they would like to be supported by tools that enable digital oblivion.

This study is not designed to show the possible consequences of such mechanisms and ask whether users would accept those in exchange for digital oblivion in OSN. For example, it is possible that a tool implementing digital oblivion needs access to all posts of a user or private information.

[1] `https://www.twitter.com`
[2] `https://www.facebook.com`
[3] `https://www.instagram.com`

5.5.5 Abstraction and Missing Implementation Details of Introduced Tools

In the survey, several digital tools were introduced to the participants. It was not described in detail what a user has to do to use these tools properly.

The tools were chosen in a way that the participants consider it plausible that those tools might exist now or in the near future. This thesis did not contribute to the question of how exactly these tools could be implemented. It might be that some of them can not be implemented at all.

6 Conclusion

This chapter first gives a summary of this thesis with an emphasis on the study results (for result details see chapter 4, for details on the discussion see chapter 5). It then provides an outlook on future work based on this thesis.

6.1 Summary

This thesis provides two contributions to the research on digital oblivion. As first contribution, it provides an overview of arguments for and against digital oblivion found in literature.

The discussion on whether digital oblivion should be implemented or not is controversial. Both sides argue that the absence or the presence of mechanisms for digital oblivion introduces a kind of censorship, restricts the freedom of expression ans speech for individuals, and is a potential danger to democracy.

Authors advocating the introduction of mechanisms for digital oblivion point out that a lack of such mechanisms leads to loss of information control for individuals, that content that can not be taken offline has the potential to ruin reputations or harm people in the future, that whatever a person does might never be forgiven by the public, and that people might start to censor themselves. Furthermore, data found online would not represent a person accurately, data might be used in unintended ways or with malicious intention, and technology could be used for surveillance.

Authors advocating against an introduction of mechanisms for digital oblivion point out that such mechanisms violate the freedom of speech, restrict access to information, can be misused for censorship, can be used to erase and rewrite history, and present a danger to democracy. When data is available and accessible, people could use it to impose pressure on companies and governments, it can aid in remembering and learning from the past, it can prove life-saving, and prevent future harm. Furthermore, comprehensive data can fuel innovation and economic growth, it can be used to create personalized content and improve the quality of life for individuals, and companies can use it to check the social qualifications of job applicants.

As the second contribution of this thesis, an anonymous user study was conducted to answer the question if the absence of forgetting mechanisms in online social networks (OSN) is a problem for users.

The survey consisted of several question groups. The central part of the survey were five scenarios taking place in OSN. Participants were asked how they would react in these scenarios, what they think, and what they consider as important in that situation. 250 participants completed the survey. About 95 % of the participants were living in German-speaking countries, mainly Germany. On average, the participants were 35 years of age.

Results show that the majority of users tries to prevent the distribution of data belonging to or being about the user when that data is shared by others. Users consider active support by the official support of the OSN as most important in this situation. Tools implementing aspects of digital oblivion are considered very important or important by the majority of users. Furthermore, such tools have the potential to change the decision of users who initially decided not to prevent the distribution of the data.

When someone else shares data about a user, the motivation to prevent the distribution of this data does not depend on the person who is sharing it. The motivation is between high and medium for all distributors.

Some users forget content they posted online in an OSN. About one-third of the participants would like to have a tool that displays old posts to them again after a while. Of those, almost 90 % want to view their old posts again to decide to either keep or delete them.

The majority of users does not want that their content in OSN is automatically deleted after a fixed time. However, while only few users would actually prefer an automated deletion, most users consider automated deletion as a good optional setting for each of their posts.

When users delete their OSN account, most of them are unsure whether or believe that their posts and conversation are still publicly available. The majority of users would like to have an automated tool that verifies if all their data linked to the deleted account is not publicly available anymore.

6.2 Future Work

This section provides an outlook on future work based on this thesis. The first part of this section presents future work evaluating the survey data. The second part points out future work based on the design limitations of the study. The third part of this section introduces future work based on the results of the study.

6.2.1 Further Evaluation of Data

The data of the user study was evaluated with regards to the research question: Is the absence of forgetting mechanisms in OSN a problem for users? This research question does not differentiate between users. The data collected in this study can be evaluated with another focus.

The number of male and female participants is almost equal. One could evaluate if male and female participants gave significantly different responses. Similarly, responses from participants being younger than 30 years and those being older than 30 years can be compared.

About 50 % of participants engage in IT, engineering, technique, or mathematics. This group of participants can be compared to those participants who do not engage in those fields professionally.

Participants were asked whether they use OSN professionally or privately. One could evaluate if professional users of OSN respond significantly different from private users.

6.2.2 Furture Work Based on Design Limitations

The study conducted for this thesis was designed to evaluate the absence of forgetting mechanisms in OSN. Most of the questions did not ask if users would accept a specific implementation of such a mechanism (also see subsection 5.5.4 and 5.5.5). For example, it is possible that a forgetting mechanism needs access to all posts of a user or private information. A next step would be to develop detailed concepts of such mechanisms or implement such a tool (also see subsection 6.2.3). These concepts or implementations should then be evaluated in another user study.

The study and evaluation of the results covered the surface of whom users want to defend against by enabling mechanisms for digital oblivion in OSN (see section 5.4). The question if and whom users want to defend themselves against could be explored in more depth in another user study.

6.2.3 Future Work Based on the Study Results

The results of the study indicate that some tools implementing digital oblivion are more important for users than others when data is shared about them in OSN (see subsection 5.1.3, 5.1.4, and 5.1.8). Tools finding each occurrence of the data and showing it to the user were considered most important. A next step would be to develop a model of such a tool and implement it. The usability and acceptance of such a tool should be tested in another user study.

This study did not ask why participants consider active support by one person or tool as more important than others (see subsection 5.1.5 and 5.1.8). An evaluation of this question could be part of a future user study.

A tool displaying old posts in OSN to the users from time to time with an easy option to delete the displayed post (see subsection 5.2.2 and 5.2.3) could be implemented and evaluated in a user study.

Another tool that is relevant for users is an optional setting to automatically delete content in OSN after a set time (see subsection 5.3.2). Several proposals have been made for the automated expiration of data (see section 1.2). These tools and protocols are not directly based on the OSN but utilize different infrastructures. An exception to this is the messaging service Snapchat[1] that implements the expiration of messages as part of its service.

Based on the study results, a tool or concept for a mechanism checking whether the users' content is still publicly available after they deleted their account could be developed (see subsection 5.3.5). If a specific tool is implemented, this tool should be evaluated in another user study.

[1] `https://www.snapchat.com`

A Survey Details

This chapter contains the survey that was presented to the participants. The survey was provided in German and English, this section only includes the English version.

Mandatory questions are marked with a * after the question. If a question is only shown under certain conditions, those conditions are written down in *italic font* and with a black diamond (⧫) before and after the condition. Additional information, like randomization of the following question groups, is written down in *italic font* and surrounded by two white diamonds (♢). A horizontal line separates different sections of the survey.

User-Sided Information Control in Social Media

Goal of the Study
This study examines the behavior of users with (shared) content within social media. We want to find out how users react in certain situations, what is important for them and what they consider as helpful in these situations.

Expected Time
The expected time for your participation is about 10 minutes.

What Happens When You Decide to Withdraw From the Study
You are free to withdraw from this study at any time without any kind of penalty.

You can quit the participation by closing the browser you opened the survey with. Answers you gave until that moment will not be used in the results of the study.

You can not withdraw from the study after you finished filling out the survey due to data being saved anonymously. This means that it is not possible for us to assign received answers to a single participant and delete those answers.

Data Protection and Consent

Collected Data
Result data (meaning the answers given on questions and tasks of the survey), date

and time of your participation as well as the time you need to fill out the survey are collected during your participation and stored for statistical purpose.

Data Protection and Usage
This survey is conducted anonymously. This means that it is impossible to draw any conclusions on who a participant is.

The collected data is used for analysis within this study. Results and data will be published as scientific publication in printed and digital form. Data might be handed over to various scientific research institutes or made publicly available on the internet.

Contact Information
This study is conducted by Maria Kober and Florian Farke at the Department of Electrical Engineering and Information Technology of the Ruhr-University Bochum. If you have questions, problems or notes about this study or your participation, please write an email to Maria Kober (maria.kober+survey@rub.de).

EXPLANATION OF CONSENT by the participant:
By clicking Next I confirm that I have read above information on participation, that I have understood the content, that all questions I had regarding this are answered and that I agree to these terms and conditions for participation in this study.

Thank you very much for your support!

Demography

How old are you?*

Please specify your gender:

- ◯ male
- ◯ female
- ◯ other
- ◯ prefer not to say

Which continent do you live on?

- ◯ Africa
- ◯ Asia
- ◯ Australia and Oceania
- ◯ Europe

◯ North-America

◯ South-America

♦ *Only countries situated on the selected continent are shown.* ♦
Which country do you live on?
♢ *A list of all countries is displayed here.* ♢

♦ *This section is only shown when participants are younger than 18 years of age.* ♦

Age-Restricted Participation

Thank you very much for your willingness to participate in this study! Unfortunately we can only allow participants who are at least 18 years of age.

General Questions on Social Media

What do the terms "social media" and "social networks" mean in this survey?
Social media and social networks here mean online services where you can create a profile and interact, communicate and exchange content with other users of the service. They all have in common that at least parts of the user activity - e.g., like sharing images - are publicly visible. Some social networks provide the option to restrict the visibility of one's own activity.

Please indicate if and how you use the following social media services:

	How often do you use the social network?					Private or business usage?			
	Daily	Several times a week	Several times a month	Scarcely	Never	Private usage	Business usage	Private and business usage	Neither
Facebook	◯	◯	◯	◯	◯	◯	◯	◯	◯
Flickr	◯	◯	◯	◯	◯	◯	◯	◯	◯
Instagram	◯	◯	◯	◯	◯	◯	◯	◯	◯
LinkedIn	◯	◯	◯	◯	◯	◯	◯	◯	◯
MySpace	◯	◯	◯	◯	◯	◯	◯	◯	◯
Pinterest	◯	◯	◯	◯	◯	◯	◯	◯	◯
Tumblr	◯	◯	◯	◯	◯	◯	◯	◯	◯
Twitter	◯	◯	◯	◯	◯	◯	◯	◯	◯
Xing	◯	◯	◯	◯	◯	◯	◯	◯	◯

If there are other social networks you are using at least once a month, which are those?

♦ *This question is displayed if the participant chose* Never *for each social network, and did not type in another social network he is using.* ♦
Why are you no active user of social media?

♦ *This question is displayed if the participant uses at least one social network.* ♦
Please estimate how you spend your time in social networks:

- ◯ I spend most of the time to maintain my own profile and to upload own content
- ◯ I spend most of the time to follow the content of other users
- ◯ I spend about the same amount of time to maintain my profile and to follow the content of other users

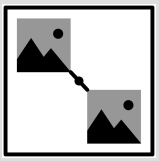
If there are social media services which you don't use anymore, what has caused your inactivity?

Scenarios

On the next pages, **5 scenarios** will be presented to you. Each scenario takes place in a social network.
If you experienced one of the situations first-hand in a social network, please answer according to your experiences. Otherwise please imagine how you would act in this situation.

♢ *The following five scenarios are shown in random order.* ♢

Scenario: Image Reference



Imagine you published an image on Facebook. Later you deleted the post including the image. After the post was deleted, you find copies of your image within the posts of other people. You do not know if someone actively distributes a screenshot of your image or if this is a reference on your deleted post still displaying the image.

If you experienced this situation first-hand in a social network, please answer according to your experiences. Otherwise please imagine how you would act in this situation.

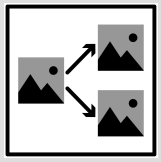
Do you consider it problematic that your image is still distributed even though you deleted it?

- ◯ Yes, in any case
- ◯ Yes, in most cases
- ◯ It strongly depends on the content
- ◯ No, I usually have no problem with that
- ◯ No, I never have a problem with that

Did you ever experience this scenario within a social network?

- ◯ Yes
- ◯ No

Scenario: Screenshot Sharing



You uploaded and shared an image on Instagram. The picture is neither embarrassing for you nor does it depict sensible information, yet you do not want this image to be shared publicly. You have taken appropriate actions to ensure this, for example you restricted the visibility or deleted the image some time later.

Now you discover that someone has taken a screenshot of your post and shares this screenshot publicly in the social network.

If you experienced this situation first-hand in a social network, please answer according to your experiences. Otherwise please imagine how you would act in this situation.

Do you actively try to prevent or restrict the distribution?*

- ◯ Yes
- ◯ No

♦ *This question is displayed if the first question was answered with* Yes. ♦
How important is the active support of the following people, groups or tools for your effort to prevent the distribution of the screenshot?

	no support wanted	not important at all	of little importance	important	very important
Family	◯	◯	◯	◯	◯
Friends	◯	◯	◯	◯	◯
Official support of the social network	◯	◯	◯	◯	◯
Follower/Fans	◯	◯	◯	◯	◯
Strangers	◯	◯	◯	◯	◯
Tools that automatically find each occurrence of the screenshot and show it to me	◯	◯	◯	◯	◯
Tools that automatically delete each occurrence of the screenshot	◯	◯	◯	◯	◯
Tools that automatically report each occurrence of the screenshot to the official support of the social network	◯	◯	◯	◯	◯

♦ *This question is displayed if the first question was answered with* Yes. ♦
Are there other people, groups or tools you would like to be actively supported by in this situation?

♦ *This question is displayed if the first question was answered with* Yes. ♦
How high is your motivation to prevent the distribution of the screenshot if one of the following people shared this screenshot?

	low	medium	high
Family members	◯	◯	◯
Friends	◯	◯	◯
Other acquaintances	◯	◯	◯
Follower/Fans	◯	◯	◯
Fellow workers	◯	◯	◯
Superiors	◯	◯	◯
Strangers	◯	◯	◯

♦ *This question is displayed if the first question was answered with* No. ♦
Why do you not try to prevent the distribution of the screenshot?

- ☐ It is not important enough for me
- ☐ I don't want to spend time on this
- ☐ I don't know how to contact the person who shared the screenshot
- ☐ I feel like the support team of the social network does not support me in a helpful way
- ☐ I do not know how to find all occurrences of the screenshot
- ☐ I feel like an action on my side would not change anything
- ☐ An action on my side would only make the situation worse
- ☐ Actually, I now appreciate that the screenshot was shared
- ☐ Other: ☐

♦ *This question is displayed if the first question was answered with* No. ♦
Would you try to prevent or restrict the distribution of the screenshot if the following people supported you actively?

- ☐ Family members
- ☐ Friends
- ☐ Follower/Fans
- ☐ Strangers

♦ *This question is displayed if the first question was answered with* No. ♦
Would assistance by the official support lead to you trying to prevent or restrict the distribution of the screenshot?

◯ Yes

◯ No

♦ *This question is displayed if the first question was answered with* No. ♦
Would support through the following tools lead to you trying to prevent or restrict the distribution of the screenshot?

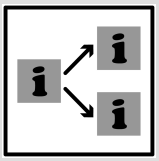
☐ Tools that automatically find each occurrence of the screenshot and show it to me

☐ Tools that automatically delete each occurrence of the screenshot

☐ Tools that automatically report every occurrence of the screenshot to the official support of the social network

---

Scenario: Information Spreading Through Third Parties



You come across a public post on Facebook where the author is sharing information about you. You do not want this information to be publicly visible on Facebook; actually, this information should never be shared on any social network.

If you experienced this situation first-hand in a social network, please answer according to your experiences. Otherwise please imagine how you would act in this situation.

Do you actively try to prevent or restrict the distribution?*

◯ Yes

◯ No

♦ *This question is displayed if the first question was answered with* Yes. ♦
How important is the active support of the following people, groups or tools for your effort to prevent the distribution of the information?

	no support wanted	not important at all	of little importance	important	very important
Family	◯	◯	◯	◯	◯
Friends	◯	◯	◯	◯	◯
Official support of the social network	◯	◯	◯	◯	◯
Follower/Fans	◯	◯	◯	◯	◯
Strangers	◯	◯	◯	◯	◯
Tools that automatically find each occurrence of the information and show it to me	◯	◯	◯	◯	◯
Tools that automatically delete each occurrence of the information	◯	◯	◯	◯	◯
Tools that automatically report each occurrence of the information to the official support of the social network	◯	◯	◯	◯	◯

♦ *This question is displayed if the first question was answered with* Yes. ♦
Are there other people, groups or tools you would like to be actively supported by in this situation?

♦ *This question is displayed if the first question was answered with* Yes. ♦
How high is your motivation to prevent the distribution of the information if one of the following people shared this information?

	low	medium	high
Family members	◯	◯	◯
Friends	◯	◯	◯
Other acquaintances	◯	◯	◯
Follower/Fans	◯	◯	◯
Fellow workers	◯	◯	◯
Superiors	◯	◯	◯
Strangers	◯	◯	◯

♦ *This question is displayed if the first question was answered with* No. ♦
Why do you not try to prevent the distribution of the screenshot?

- ☐ It is not important enough for me
- ☐ I don't want to spend time on this
- ☐ I don't know how to contact the person who shared the information
- ☐ I feel like the support team of the social network does not support me in a helpful way
- ☐ I do not know how to find all occurrences of the information
- ☐ I feel like an action on my side would not change anything
- ☐ An action on my side would only make the situation worse
- ☐ Actually, I now appreciate that the information was shared
- ☐ Other: ☐

♦ *This question is displayed if the first question was answered with* No. ♦
Would you try to prevent or restrict the distribution of the information if the following people supported you actively?

- ☐ Family members
- ☐ Friends
- ☐ Follower/Fans
- ☐ Strangers

♦ *This question is displayed if the first question was answered with* No. ♦
Would assistance by the official support lead to you trying to prevent or restrict the distribution of the information?

◯ Yes

◯ No

♦ *This question is displayed if the first question was answered with* No. ♦
Would support through the following tools lead to you trying to prevent or restrict the distribution of the screenshot?

☐ Tools that automatically find each occurrence of the information and show it to me

☐ Tools that automatically delete each occurrence of the information

☐ Tools that automatically report every occurrence of the information to the official support of the social network

In case you experienced this scenario within a social network, how did you get to know about the post?

☐ I have not yet experienced this situation

☐ I got to know about this post on coincidence

☐ The post was directly addressed at myself

☐ I got to know about this post through followers or fans

☐ I got to know about this post through friends, family or acquaintances

☐ I got to know about this post through a stranger

☐ I got to know about the post through other means: ☐

Scenario: Account Deletion

You published several tweets with your Twitter account and interacted with other members of the social network.

Now you delete your Twitter account.

If you experienced this situation first-hand in a social network, please answer according to your experiences. Otherwise please imagine how you would act in this situation.

Do you think that some of your tweets and conversations are still publicly available after the deletion?

◯ Yes

◯ No

◯ I am unsure

♦ *This question is displayed if the previous question was answered with* Yes *or* I am unsure. ♦
Would you consider it problematic if your content and conversations were still publicly visible after the deletion?

◯ Yes

◯ No

Would you like to have an app or a digital tool that verifies that all occurrences of your tweets and conversations on Twitter are not publicly available anymore?

◯ Yes

◯ No

Do you think that Twitter deletes your tweets and conversations from their servers shortly after you deleted your account?

◯ Yes

◯ No

◯ I am unsure

♦ *This question is displayed if the previous question was answered with* No *or* I am unsure. ♦
Would you consider it problematic if your content and conversations were stored long-term on Twitter servers?

◯ Yes

◯ No

◯ I am unsure

Scenario: Forgotten Image

On Instagram, one of your fans/followers asks you about an image. You vaguely recognize the image but you don't recall where you know it from. A short research shows: you uploaded this image on Instagram some time ago and forgot about this post over time.

If you experienced this situation first-hand in a social network, please answer according to your experiences. Otherwise please imagine how you would act in this situation.

Did you experience this scenario within a social network?

◯ Yes

◯ No

Would you like to have a setting in your account so that images you uploaded will be shown to you again after a certain period of time?

◯ Yes

◯ No

♦ *This question is displayed if the previous question was answered with* Yes. ♦
Why do you want old posts to be displayed to you again?

☐ So that I don't forget them

☐ So that I look at my old pictures more often

☐ So that I can decide to keep or delete them

☐ So that I can revise my post after a while

☐ Due to other reasons: ____

Do you want your posts to be deleted automatically after a set period of time?

◯ Yes

◯ Yes, and I want to set the time until the posts are deleted by myself

◯ No, but this sounds good as an extra option for each post

◯ No

Demography

What is the highest level of school you have completed or the highest degree you have received?

- ◯ No school degree
- ◯ Less than high school degree or equivalent
- ◯ Finished vocational training
- ◯ School leaving qualification
- ◯ Bachelor degree
- ◯ Master degree/graduate degree
- ◯ Doctor degree
- ◯ Other: ☐

Which field describes your current activity best?

In case you are studying at university/college or are in vocational training, please choose the subject of your study/training.

- ◯ Administration, Management, Law
- ◯ Art, Culture, Literature
- ◯ Craftsmen
- ◯ Economic Sciences
- ◯ Education, Social
- ◯ Housewife or houseman
- ◯ IT, Engineering, Technique, Mathematics
- ◯ Media, Communication, Advertisement
- ◯ Medicine, Health, Psychology
- ◯ Natural Sciences, Life Sciences
- ◯ Other
- ◯ Police, Military, Personal Security
- ◯ Politics, Political Sciences
- ◯ Prefer not to say

Do you engage privately or professionally in:

	Professionally			Privately		
	Yes	A little bit	No	Yes	A little bit	No
Social Media	◯	◯	◯	◯	◯	◯
Data Privacy and Data Protection	◯	◯	◯	◯	◯	◯
IT-Security	◯	◯	◯	◯	◯	◯

Thank you very much for your time and your participation!

You can close the browser window now.

In case you have any questions or notes about this study or your participation, please write an email to Maria Kober (maria.kober+survey@rub.de).

B Participant Recruitment Texts

This chapter contains texts that were written to recruit participants and make sharing of the survey link easier.

B.1 For Mailing Lists

The following text was written for the mailing lists of Ruhr University Bochum[1]. The main target were students of various faculties.

Liebe Kommilitonen/Dear fellow students,
— English Version below —

in meiner Abschlussarbeit beschäftige ich mich mit der Frage, wie man als Nutzer von sozialen Medien die Kontrolle über seine eigenen Informationen und Daten behält. Dazu brauche ich eure Unterstützung: meine Forschung basiert auf einem 10-minütigen Fragebogen, der die Teilnehmer (also hoffentlich auch dich) zu verschiedenen Situationen in sozialen Medien befragt. Auch wenn du keine sozialen Medien nutzt, bist du herzlich eingeladen, an der Umfrage teilzunehmen: `https://mobsec-studies.rub.de/index.php?r=survey/index&sid=437948&lang=de`

In my master's thesis I study the question how users of social media can retain control over their own information and data. For this, I do need your help: my research is based on a 10-minute survey in which I ask the participants (so hopefully you, too) about different situations in social media. Even if you do not use social media, you are very welcome to participate in the survey: `https://mobsec-studies.rub.de/index.php?r=survey/index&sid=437948&lang=en`

Vielen Dank/Thank you very much,
Maria Kober

P.s. Den Umfragelink und diese Beschreibung kannst du gerne mit anderen teilen.
P.s. Please feel free to share this survey link and description with others.

[1] `https://lists.ruhr-uni-bochum.de/mailman/listinfo`

B.2 For Social Media

The following text was written for use within social media and it was sometimes adapted to the environment it was posted in. It was the most distributed information text to recruit participants.

Social media and control over one's own information? This is what I study within the scope of my final thesis at Ruhr-University Bochum. To advance this research, I need your support by participating in a short, anonymous survey: `https://mobsec-studies.rub.de/index.php?r=survey/index&sid=437948&lang=en`
Even if you do not use social media, you are very welcome to participate in the survey.
P.s. Please feel free to share this survey link and description with others.

C Survey – Additional Results

This chapter presents additional results to the questions why participants are inactive users of online social networks (OSN).

Participants were asked why they are no active users of OSN. 15 participants wrote down an answer to this question. Those answers were sorted into one or more categories. Table C.1 presents the categories and the number of participants who specified this as reason to not use certain OSN anymore. The two most-named reasons are that the participants have no need for OSN and that they consider them a waste of time.

Table C.1: Participants were asked why they are no active users of OSN. 15 participants responded to this question.

	Participants
No need for OSN	8
Waste of time	5
Privacy issues	4
Prefer personal contact	4
Excessive demands	2

Participants were asked if and why they stopped using some OSN. 84 participants wrote down an answer to this question. Those answers were sorted into one or more categories. Table C.2 presents the categories and the number of participants who specified this as reason to not use certain OSN anymore. The main reasons are that the participants have no need for the OSN anymore, that the content became uninteresting, annoying, or toxic, and that users have privacy concerns.

Table C.2: Participants were asked if and why they stopped using some OSN. 84 participants responded to this question.

	Participants
No further need	22
Uninteresting, annoying, or toxic	22
Privacy issues	19
Too time-consuming	16
Usability	8
Friends left the OSN	6
Unsuitable for personal contact	6
Loss of trust	2
Other	4

D Acronyms

OSN online social networks

RUB Ruhr University Bochum

List of Figures

List of Tables

Bibliography

[1] Liam J. Bannon. "Forgetting As a Feature, Not a Bug: The Duality of Memory and Implications for Ubiquitous Computing". In: *CoDesign* 2.1 (Mar. 2006), pp. 3–15.

[2] Matt Bishop, Emily Rine Butler, Kevin Butler, Carrie Gates, and Steven Greenspan. "Forgive and Forget: Return to Obscurity". In: *New Security Paradigms Workshop*. NSPW '13. New York, New York, USA: ACM, Sept. 2013, pp. 1–10.

[3] Viktor Mayer-Schönberger. *Delete: The Virtue of Forgetting in the Digital Age*. First Edition. Princeton, New Jersey, USA: Princeton University Press, July 2009.

[4] Martin Hilbert and Priscila López. "The World's Technological Capacity to Store, Communicate, and Compute Information". In: *Science* 332.6025 (2011), pp. 60–65.

[5] Kimio Tatsuno. *Current Trends in Digital Cameras and Camera-Phones*. Tech. rep. NISTEP Science & Technology Foresight Center, 2006.

[6] Peter Cohen. *A History of Hard Drives*. `https://www.backblaze.com/blog/history-hard-drives/`, as of March 9, 2019. 2016.

[7] DOMO. *Data Never Sleeps 5.0*. `https://www.domo.com/learn/data-never-sleeps-5`, as of March 9, 2019. 2017.

[8] Simon Kemp. *Digital in 2017 Global Overview – a collection of internet, social media, and mobile data from around the world*. Published by We Are Social and Hootsuite. `https://www.slideshare.net/wearesocialsg/digital-in-2017-global-overview`, as of December 15, 2018. 2017.

[9] DOMO. *Data Never Sleeps 6.0*. `https://www.domo.com/learn/data-never-sleeps-6`, as of December 15, 2018. 2018.

[10] Alexander Novotny and Sarah Spiekermann. "Oblivion on the Web: An Inquiry of User Needs and Technologies". In: *European Conference on Information Systems*. ECIS '14. Tel Aviv, Israel: Association for Information Systems, June 2014.

[11] Lujo Bauer, Lorrie Faith Cranor, Saranga Komanduri, Michelle L. Mazurek, Michael K. Reiter, Manya Sleeper, and Blase Ur. "The Post Anachronism: The Temporal Dimension of Facebook Privacy". In: *ACM Workshop on Privacy in the Electronic Society*. WPES '13. New York, New York, USA: ACM, 2013, pp. 1–12.

[12] Ambar Murillo, Andreas Kramm, Sebastian Schnorf, and Alexander De Luca. ""If I press delete, it's gone" - User Understanding of Online Data Deletion and Expiration". In: *Symposium on Usable Privacy and Security*. SOUPS '18. Baltimore, Maryland, USA: USENIX, Aug. 2018, pp. 329–339.

[13] Mainack Mondal, Johnnatan Messias, Saptarshi Ghosh, Krishna P. Gummadi, and Aniket Kate. "Longitudinal Privacy Management in Social Media: The Need for Better Controls". In: *IEEE Internet Computing* 21.3 (May 2017), pp. 48–55.

[14] Oshrat Ayalon and Eran Toch. "Retrospective Privacy: Managing Longitudinal Privacy in Online Social Networks". In: *Symposium on Usable Privacy and Security*. SOUPS '13. Newcastle, United Kingdom: ACM, July 2013, 4:1–4:13.

[15] Oshrat Ayalon and Eran Toch. "Not Even Past: Information Aging and Temporal Privacy in Online Social Networks". In: *Human-Computer Interaction* 32.2 (July 2016), pp. 73–102.

[16] Kurt Thomas, Chris Grier, and David M. Nicol. "unFriendly: Multi-Party Privacy Risks in Social Networks". In: *International Symposium on Privacy Enhancing Technologies Symposium*. 2010, pp. 236–252.

[17] Ieng-Fat Lam, Kuan-Ta Chen, and Ling-Jyh Chen. "Involuntary Information Leakage in Social Network Services". In: *International Workshop on Security*. Springer. 2008, pp. 167–183.

[18] Andrew Besmer and Heather Richter Lipford. "Moving Beyond Untagging: Photo Privacy in a Tagged World". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2010, pp. 1563–1572.

[19] Mary Madden. "Privacy Management on Social Media Sites". In: *Pew Internet Report* (2012), pp. 1–20.

[20] Mary Madden and Aaron Smith. *Reputation Management and Social Media*. Retrieved from the Pew Research Center `https://www.pewinternet.org/2010/05/26/reputation-management-and-social-media/`. 2010.

[21] Airi Lampinen, Vilma Lehtinen, Asko Lehmuskallio, and Sakari Tamminen. "We're in It Together: Interpersonal Management of Disclosure in Social Network Services". In: *Proceedings of the SIGCHI conference on human factors in computing systems*. ACM. 2011, pp. 3217–3226.

[22] Radia Perlman. *The Ephemerizer: Making Data Disappear*. Technical Report SMLI TR-2005-140. Sun Microsystems Laboratories, Feb. 2005.

[23] Roxana Geambasu, Tadayohsi Kohno, Amit A. Levy, and Henry M. Levy. "Vanish: Increasing Data Privacy with Self-Destructing Data". In: *USENIX Security Symposium*. SSYM '09. Montreal, Quebec, Canada: USENIX, Aug. 2009, pp. 299–316.

[24] Lingfang Zeng, Zhan Shi, Shengjie Xu, and Dan Feng. "SafeVanish: An Improved Data Self-Destruction for Protecting Data Privacy". In: *IEEE Conference on Cloud Computing Technology and Science*. CloudCom '10. Indianapolis, Indiana, USA: IEEE Computer Society, Nov. 2010, pp. 521–528.

[25] Julian Backes, Michael Backes, Markus Dürmuth, Sebastian Gerling, and Stefan Lorenz. "X-pire! - A Digital Expiration Date for Images in Social Networks". In: *CoRR* abs/1112.2649 (Dec. 2011), pp. 1–22.

[26] Claude Castelluccia, Emiliano De Cristofaro, Aurélien Francillon, and Mohamed-Ali Kaafar. "EphPub: Toward Robust Ephemeral Publishing". In: *IEEE Conference on Network Protocols*. ICNP '11. Vancouver, British Columbia, Canada: IEEE Computer Society, Oct. 2011, pp. 165–175.

[27] Sirke Reimann and Markus Dürmuth. "Timed Revocation of User Data: Long Expiration Times from Existing Infrastructure". In: *ACM Workshop on Privacy in the Electronic Society*. WPES '12. Raleigh, North Carolina, USA: ACM, Oct. 2012, pp. 65–74.

[28] Apostolis Zarras, Katharina Kohls, Markus Dürmuth, and Christina Pöpper. "Neuralyzer: Flexible Expiration Times for the Revocation of Online Data". In: *ACM Conference on Data and Application Security and Privacy*. CODASPY '16. New Orleans, Louisiana, USA: ACM, Mar. 2016, pp. 14–25.

[29] Snapchat Support. *When does Snapchat delete Snaps and Chats?* Online. `https://support.snapchat.com/en-US/article/when-are-snaps-chats-deleted` as of October 13, 2019. 2019.

[30] Jongwon Lee. "What the Right to be Forgotten Means to Companies: Threat or Opportunity?" In: *Procedia Computer Science* 91 (July 2016), pp. 542–546.

[31] Chanhee Kwak, Junyeong Lee, and Heeseok Lee. "Forming a Dimension of Digital Human Rights: Research Agenda for the Right to be Forgotten". In: *Hawaii International Conference on System Sciences*. HICSS '17. Hilton Waikoloa Village, Hawaii: AIS, Jan. 2017.

[32] Ghous Amjad, Muhammad Shujaat Mirza, and Christina Pöpper. "Forgetting with Puzzles: Using Cryptographic Puzzles to support Digital Forgetting". In: *ACM Conference on Data and Application Security and Privacy*. CODASPY '18. Tempe, Arizona, USA: ACM, Mar. 2018, pp. 342–353.

[33] Chris Conley. "The Right to Delete". In: *AAAI Spring Symposium: Intelligent Information Privacy Management*. Palo Alto, California, USA: Stanford University, Mar. 2010.

[34] Olga Kieselmann, Nils Kopal, and Arno Wacker. “A Novel Approach to Data Revocation on the Internet”. In: *International Workshop on Data Privacy Management, and Security Assurance*. QASA ’15. Vienna, Austria: Springer, Sept. 2016, pp. 134–149.

[35] Martha Garcia-Murillo and Ian MacInnes. *The Right to Be Forgotten: Its Weaknesses and Alternatives*. Nov. 2014.

[36] Noah Hirsch, Chris Kanich, Mohammad Taha Khan, Xuefeng Liu, Mainack Mondal, Michael Tang, Christopher Tran, Blase Ur, William Wang, Günce Su Yilmaz, and Elena Zheleva. “Making Retrospective Data Management Usable”. In: *Proceedings of the Fourteenth Symposium on Usable Privacy and Security (SOUPS)*. Baltimore, Maryland, USA: ACM, Aug. 2018.

[37] Daniel J. Solove. “The Future of Reputation: Gossip, Rumor, and Privacy on the Internet”. In: *GWU Law School Public Law Research Paper 2017-4* (2007).

[38] Jeffrey Rosen. “Free Speech, Privacy, and the Web That Never Forgets”. In: *Journal on Telecommunications and High Technology Law* 9 (2011), p. 345.

[39] Meg Leta Ambrose and Jef Ausloos. “The Right to Be Forgotten Across the Pond”. In: *Journal of Information Policy* 3 (2013), pp. 1–23.

[40] Leticia Bode and Meg Leta Jones. “Ready to Forget: American Attitudes Toward the Right to be Forgotten”. In: *The Information Society* 33.2 (Sept. 2017), pp. 76–85.

[41] Hee Joo Lee, Jang Ho Yun, Hyun Sik Yoon, and Kyung Ho Lee. “The Right to be Forgotten: Standard on Deleting the Exposed Personal Information on the Internet”. In: *Computer Science and its Applications*. CSA ’15. Berlin, Germany: Springer, Dec. 2015, pp. 883–889.

[42] Yuriko Haga. “Right to be Forgotten: A New Privacy Right in the Era of Internet”. In: *New Technology, Big Data and the Law*. First. Singapore, Philippines: Springer, Sept. 2017. Chap. 4, pp. 97–126.

[43] Shuzhe Yang. “Why Are People so Naïve? Long-term Motivation in Online Reputation Management: A Grounded Theory Study”. In: *Americas Conference on Information Systems*. AMCIS ’15. Fajardo, Puerto Rico: AIS, June 2015.

[44] Jeffrey Rosen. “The Web Means the End of Forgetting”. In: *The New York Times* 21 (2010).

[45] Anita L. Allen. “Dredging up the Past: Lifelogging, Memory, and Surveillance”. In: *University of Chicago Law Review* 75 (2008), pp. 47–74.

[46] Norberto Nuno Gomes de Andrade. “Oblivion: The Right to be Different ... from Oneself: Re-Proposing the Right to be Forgotten”. In: *The Ethics of Memory in a Digital Age: Interrogating the Right to be Forgotten*. First Edition. London, United Kingdom: Palgrave Macmillan, Jan. 2014. Chap. 5, pp. 65–81.

[47] Meg Leta Ambrose. “It’s About Time: Privacy, Information Life Cycles, and the Right to be Forgotten”. In: *Stanford Technology Law Review* 16.2 (Jan. 2013).

[48] Jonathan Zittrain. *The Future of the Internet–And How to Stop It*. Yale University Press, 2008.

[49] Paulan Korenhof, Jef Ausloos, Iva Székely, Megn Ambrose, Ronald Leenes, and Giovanni Sartor. “Timing the Right to Be Forgotten: A Study into ’Time’ as a Factor in Deciding About Retention or Erasure of Data”. In: *Reforming European Data Protection Law*. First Edition. Dordrecht, Netherlands: Springer, Oct. 2015. Chap. 7, pp. 171–201.

[50] Manya Sleeper, Rebecca Balebako, Sauvik Das, Amber Lynn McConahy, Jason Wiese, and Lorrie Faith Cranor. “The Post that Wasn’t: Exploring Self-Censorship on Facebook”. In: *Proceedings of the 2013 conference on Computer supported cooperative work*. ACM, 2013, pp. 793–802.

[51] Jean-François Blanchette and Deborah G. Johnson. “Data Retention and the Panoptic Society: The Social Benefits of Forgetfulness”. In: *The Information Society* 18.1 (2002), pp. 33–45.

[52] Alexander Novotny. “Signs of Time: Designing Social Networking Site Profile Interfaces with Temporal Contextual Integrity”. In: *Human Aspects of Information Security, Privacy, and Trust*. HAS ’15. Los Angeles, California, USA: Springer, July 2015, pp. 547–558.

[53] Rolf H. Weber. “The Right to be Forgotten: More Than a Pandora’s Box”. In: *Journal of Intellectual Property, Information Technology and Electronic Commerce Law* 2.2 (July 2011), pp. 120–130.

[54] Jack Herlocker. *Stacy Snyder and the Untruth That Won’t Die*. `https://medium.com/@jherlocker/stacy-snyder-and-the-untruth-that-won-t-die-549c2f525661`, as of March 13, 2019. Sept. 2015.

[55] Julian Sanchez. *Court rejects appeal over student-teacher drunk MySpace pics*. `https://arstechnica.com/tech-policy/2008/12/court-rejects-appeal-over-student-teacher-drunk-myspace-pics/`, as of April 13, 2019. Dec. 2008.

[56] Armen Hareyan. *Millersville University Statement on Stacy Snyder Lawsuit*. `http://www.huliq.com/20467/millersville-university-statement-on-stacy-snyder-lawsuit`, as of May 22, 2019. May 2007.

[57] Andrew Levy. *Teenage office worker sacked for moaning on Facebook about her ’totally boring’ job*. Published by The Daily Mail. `https://www.dailymail.co.uk/news/article-1155971/Teenage-office-worker-sacked-moaning-Facebook-totally-boring-job.html`, as of February 25, 2019. 2009.

[58] Milivoj Simeonovski, Fabian Bendun, Muhammad Rizwan Asghar, Michael Backes, Ninja Marnau, and Peter Druschel. "Oblivion: Mitigating Privacy Leaks by Controlling the Discoverability of Online Information". In: *Applied Cryptography and Network Security*. ACNS '13. New York, New York, USA: Springer, June 2015, pp. 431–453.

[59] Michael Douglas. "Questioning the Right to Be Forgotten". In: *Alternative Law Journal* 40.2 (June 2015), pp. 109–112.

[60] Anna Bunn. "The Curious Case of the Right to be Forgotten". In: *Computer Law & Security Review* 31.3 (Apr. 2015), pp. 336–350.

[61] European Parliament and Council of the European Union. "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)". In: *Official Journal of the European Union* 119.1 (Apr. 2016), pp. 1–88.

[62] Isaac Arnsdorf. *Seattle attorney finds that the Internet won't let go of his past*. Published by The Seattle Times. `https://www.seattletimes.com/seattle-news/seattle-attorney-finds-that-the-internet-wont-let-go-of-his-past/`, as of February 25, 2019. 2008.

[63] Charles Arthur. *Wikipedia sued by German killers in privacy claim*. Online. `https://www.theguardian.com/technology/2009/nov/13/wikipedia-sued-privacy-claim` as of August 24, 2019. Nov. 2009.

[64] Jeffrey Rosen. "The Right to Be Forgotten". In: *Stanford Law Review Online* 64.88 (Feb. 2011).

[65] Klara Stokes and Niklas Carlsson. "A Peer-to-Peer Agent Community for Digital Oblivion in Online Social Networks". In: *Conference on Privacy, Security and Trust*. PST '13. Tarragona, Spain: IEEE, July 2013, pp. 103–110.

[66] Peter Fleischer. *Foggy thinking about the Right to Oblivion*. Online. `https://peterfleischer.blogspot.com/2011/03/foggy-thinking-about-right-to-oblivion.html` as of August 18, 2019. 2011.

[67] Kent Walker. *Defending access to lawful information at Europe's highest court*. `https://www.blog.google/topics/google-europe/defending-access-lawful-information-europes-highest-court/`, as of August 25, 2019. Nov. 2017.

[68] Jeroen Van Den Hoven. "Information Technology, Privacy, and the Protection of Personal Data". In: *Information Technology and Moral Philosophy* (2008), pp. 301–321.

[69] Duden. *Social Media, die*. Online. `https://www.duden.de/rechtschreibung/Social_Media` as of August 3, 2019. 2019.

[70] Duden. *Social Network, das*. Online. `https://www.duden.de/rechtschreibung/Social_Network` as of August 3, 2019. 2019.

[71] Danah M. Boyd and Nicole B. Ellison. "Social Network Sites: Definition, History, and Scholarship". In: *Journal of Computer-Mediated Communication* 13.1 (2007), pp. 210–230.

[72] Björn Rasch, Malte Friese, Wilhelm Hofmann, and Ewald Naumann. *Quantitative Methoden 1*. Second Edition. Springer Medizin Verlag Heidelberg, 2006.

[73] Stefanie Pötzsch. "Privacy Awareness: A Means to Solve the Privacy Paradox?" In: *IFIP Summer School on the Future of Identity in the Information Society*. Springer, 2008, pp. 226–236.

[74] Susan Athey, Christian Catalini, and Catherine Tucker. *The Digital Privacy Paradox: Small Money, Small Costs, Small Talk*. Working Paper. National Bureau of Economic Research, June 2017.

[75] Alex Braunstein, Laura Granka, and Jessica Staddon. "Indirect Content Privacy Surveys: Measuring Privacy without Asking about It". In: *Symposium on Usable Privacy and Security*. SOUPS '11. Pittsburgh, Pennsylvania, USA: ACM, July 2011.