INVESTIGATING DECAY REPRESENTATIONS FOR PRIVACY AND ONLINE REPUTATION MANAGEMENT

by
Riham Mohamed

A thesis submitted to
the Faculty of Graduate and Postdoctoral Affairs
in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY

Computer Science

at

CARLETON UNIVERSITY

Ottawa, Ontario
September, 2019

Abstract

The main goal of this thesis is to explore how Online Social Networks (OSNs) can support users in managing their online reputation during important selection processes. More specifically, we look at mechanisms that help social media users automatically dissociate from their past online activities, primarily in a hiring context.

First, we explore how online reputation impacts social media users through an online survey with 459 participants. We compare how the online reputation of individuals is perceived by other online users in two different contexts (employment and political). We found that online reputation does influence users' perception in both contexts, but especially when considering a *political* candidate.

One approach to helping individuals handle this issue is having older social media content gradually degrade visually until it is no longer perceivable. We conduct a lab study with 30 participants to investigate three such *decay representations* that can be applied in social media to enable users to dissociate from their past online content. We identify which *representations* match users' metaphor of *aging/decaying* and identify users' attitudes and concerns towards the concept of *aging/decaying of digital artifacts*.

Third, through an online survey with 360 managers, we evaluate how a decay representation impacts their assessment of users' online reputation. We compared managers' reaction to a decayed profile for a fictitious job candidate with their reaction to an original profile or an empty profile. We found that the decay representation led to significantly more positive hiring decisions and assessments of the candidate.

Fourth, we further evaluated the use of a *decay representation* on social media profiles through in-person interviews with 48 managers. Once again, the decay representation positively influenced managers' assessments, and managers discussed active evaluation of social media profiles in their real-life hiring decisions. Results confirmed that online reputation influences hiring practices, further emphasizing the need for tools that allow users to automatically dissociate from obsolete online content.

Finally, we discuss practical aspects related to decay representations and propose preliminary recommendations for the privacy and HCI communities.

List of Acronyms

EU European Union.

GDPR General Data Protection Regulation.

HCI Human-Computer Interaction.

HR Human Resources.

OPC Office of the Privacy Commissioner of Canada.

ORM Online Reputation Management.

OSN Online Social Network.

PbD Privacy by Design.

RTBF Right to Be Forgotten.

UI User Interface.

Acknowledgements

I would like to begin by thanking my thesis supervisor, Sonia Chiasson, for her continuous support. She has been a great example and a role model to me, on both a personal and a professional level. She always made herself available and provided unlimited advice. As someone with a high level of knowledge and a broad spectrum of expertise, Prof. Chiasson has given me invaluable guidance throughout the duration of my degree and I am deeply indebted to her.

I would also like to thank my defence committee, Pamela Briggs, Daniel Amyot, Audrey Girouard, and Andrew Patrick. They have provided me with useful guidance and valuable feedback that helped inform various parts of the thesis.

I am also grateful to my parents and my brothers. I would like to thank them for believing in me and for always encouraging me to achieve my goals. Although they are so far away, they never ceased to give me the emotional support and encouragement I needed. I would not have been able to complete this work without their support.

I thank Thaís Bardini Idalino for being dedicated in our collaborative study. I am happy to have worked with such a smart woman. I also thank Khadija Baig for being such a dedicated researcher. I am grateful for her effort and time spent on transcribing interviews in one of the Chapters, and for her emotional support. Similarly, I am grateful to Paulina Chametka for her help in various tasks, including running interviews, audio transcription, and data analysis. I appreciate her time and effort.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Today, it is almost impossible for individuals to exist without having an online presence or without leaving digital traces. Most online users regularly share information on Online Social Network (OSN) sites [8, 24, 125, 127]. There are obvious personal benefits to online sharing. For example, it contributes to individuals' well-being and social interactions [1, 14, 20, 74]. However, once information is published, it is hard to control who can access, copy, or modify it. Moreover, the private versus public boundaries of social media spaces are unclear, thus posting online can have unintended consequences [8]. Some consequences can present benefits to society [**?**, 112], but in many cases, consequences can be undesirable [7]. For example, employers use online data to inform hiring decisions, which might result in individuals' professional future being compromised [3, 28, 42, 48, 73, 87, 119, 123, 125, 138].

For just over a decade, employers have been using OSN data to inform their hiring decisions, in part as an inexpensive source for unofficial background checks [28, 63, 125, 133]. Up to 91% [138] of employers search OSN data during the hiring process and 70% of US recruiting professionals have rejected job candidates because of their online data [89, 127].

There are benefits to searching OSNs during the hiring process and recruitment process [69, 133]. For example, job postings on OSNs attract potential candidates and help firms connect with the right ones [125, 138]. On the other hand, this practice has shaped hiring decisions in a way that can be perceived as harmful to society [28]; current employers can be fired and new candidates not hired, sometimes as a result of relatively minor incidents or irrelevant posts [3, 28, 42, 73, 87, 119, 123, 125, 138], or sometimes indirectly based on things that are illegal to use for hiring purposes (e.g., gender [66]). Searching OSN data by employers is lacking well-grounded legal [33, 63, 125, 127, 138] or ethical [125, 133]

guidelines. Furthermore, it infringes on individuals' right to privacy, which is a main reason for disapproving of this practice.

The lines between what is private and what is relevant to professional life is also blurred with the existence of OSNs [28, 63, 127, 133]. Once information is disclosed or published to others, it is unclear what can be deemed private [28]. Moreover, users may have misconceptions about the availability of their OSN data and who can access it [127].

Meanwhile, existing privacy mechanisms and Online Reputation Management (ORM) tools do not adequately support user needs [145, 147]. In the European Union (EU), the *Right to be forgotten (RTBF)* entitles individuals, after a certain time has passed and under other specific conditions, to ask search engine companies to de-index and delete potentially damaging personal digital material. In addition, Office of the Privacy Commissioner in Canada (OPC) has been taking serious steps to investigate the potential of the RTBF under Canadian law [104] and identify solutions that balance freedom of expression and the online reputation of individuals [103]. Since the proposal of the RTBF, forgetting digital memories [87] has become an important principle to diminish the potential negative repercussions resulting from the persistent reproduction of our digital footprints. While there exists a general emphasis on reminiscing [39, 109], forgetting digital memories introduces a converse emphasis on dissociating from obsolete and irrelevant digital artifacts.

In this regard, HCI and privacy researchers have been investigating mechanisms that enable digital forgetting for users, such as content deletion [10, 30], archival [5, 10, 51], expiration [10, 101], or decay [10]. However, these mechanisms require active user engagement, which rarely happens. Alternatively, there has been interest in representing the passage of time [79] to preserve the temporal contextual integrity of previously published information [5, 16, 94, 98]. One approach visualizes the passage of time within the User Interface (UI) by using decay representations [98].

1.2 Research Statement

The goal of this research is to explore how OSNs can support users in managing their online reputation during important selection processes. More specifically, we look at mechanisms that help social media users automatically dissociate from their past online activities, primarily in a hiring context.

The objectives of this research are:

Objective 1: Understand how online reputation impacts social media users in the context of important selection processes such as a hiring process.

Objective 2: Compare different decay representations that can help social media users automatically dissociate from their past online content.

Objective 3: Identify users' attitudes and concerns towards decay representations that were presented to them.

Objective 4: Evaluate how decay representations impact users' online reputation in the context of important selection processes.

1.3 Contributions

The main contributions of the research are:

1. In Chapter 3, we explored users' perception of various online activities and whether *online reputation* should be considered by *potential employers* and the *voting public*. We launched two parallel versions of an online crowdsourcing survey. One asked questions relating to evaluating job candidates/employees based on their online presence, and the second asked the same questions of political candidates. We collected valid responses from 459 participants. Our results suggested that political candidates are held to higher standards than job candidates/employees. This work also partially supported the findings from previous work [28, 73, 87, 119, 125, 138] showing that many people have had their career compromised by their online activities and online footprint. The survey was the first to compare how the people's opinion differs depending on whether the candidate is vying for a job or political office.

2. In Chapter 4, we examined the concept of *aging/decaying of social media digital artifacts* from the user's perspective. Our goal was to identify representations that match users' metaphor of aging/decaying and explores the representations that support their privacy. Through a lab study with 30 participants, we compared three different representations that decay OSN content visible to other users on three different social media platforms. We found that *content shrinking* and *content fading* best represent aging/decaying of digital artifacts among those tested. We identified participants' attitudes and concerns towards digital aging/decaying, and discussed their preferences regarding the concept. These findings informed the design of our studies in Chapters 5 and 6 and helped us frame recommendations proposed in Chapter 7.

3. In Chapter 5, we evaluated the use of *decay representations* for ORM. We explored how *shrinking representation* applied on a job candidate's profile might affect managers' hiring decisions of the candidate. We compared with a profile showing all posts *full size* and a third profile with *no posts*. We examined whether gender of the candidate or managers' gender or age impact their decisions. We further probed managers' use of online reputation in real-life practices. We conducted a 3×2 between-subjects online survey with 360 managers where we showed them fictitious social media profiles for job candidates. Gender of the candidate had no impact and managers' age and gender had a limited impact on the results. *The shrinking representation* led to significantly more positive decisions and perception of the job candidates compared to the two tested representations. We further reported our managers' real-life practices which showed that online reputation is influential, reinforcing a need for tools that help users dissociate from their past online activities.

4. In Chapter 6, we further extended our study from Chapter 5 through a 3×2 between-subjects lab study. Using the same prototypes, questionnaires, and an additional interview guide, we recruited 48 participants with job hiring experience. We showed how the *shrinking profile* and the profile with *full size* posts better influenced participants' assessments compared to the profile with *no posts*. Moreover, the *shrinking profile* helped participants give the candidate the benefit of the doubt compared to the other two tested representations. We found that participants were more accepting of

the male candidate's negative content when shown in full size compared to the female candidate with the same content, but there was a limited impact of *participants'* age and gender. We explored our participants' use of online reputation in their practices which showed that they were more conservative compared to managers from Chapter 5. Nevertheless, online reputation could disqualify participants' prospect job candidates, confirming a need for tools that automatically dissociate user from obsolete online content. These findings as well as those from Chapter 5 further informed our recommendations in Chapter 7.

5. In Chapter 7, extending the literature and reflecting on our results, we proposed recommendations for the HCI and usable privacy communities. We discussed the potential benefit of the use of decay representations as a tool that automatically manages online reputation. We further discussed aspects related to the *feasibility* of implementing decay representations and of incorporating them into *OSN platforms' business model*, and the *social acceptability* of these representations.

1.4 Related Publications

A large portion of this thesis has been accepted or submitted to peer-reviewed academic conferences. Reham Mohamed is the first author on all of these papers. One work involved another co-author, Thaís Bardini, who is a PhD student at the University of Ottawa. A significant portion of the text appearing in this thesis is taken directly from these publications.

Full papers to be submitted are:

- **Reham Ebada Mohamed**[1] and Sonia Chiasson. [Conference Paper] Hire Me, I Have an Awesome Facebook Profile: The Influence of Decay Representations on Hiring Decisions. To be submitted to the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), 2020.

The peer-reviewed full-paper publications are:

[1]The name on the thesis is 'Riham', but the author's preferred name is 'Reham'. All publications had the name 'Reham' instead of 'Riham'.

- **Reham Ebada Mohamed** and Sonia Chiasson. [Conference Paper] Online Privacy and Aging of Digital Artifacts. In the Symposium on Usable Privacy and Security (SOUPS), USENIX Association, 2018, 12 pages.

- **Reham Ebada Mohamed**, Thaís Bardini Idalino, and Sonia Chiasson. [Conference Paper] When private and professional lives meet: The impact of digital footprints on employees and political candidates. In Social Media and Society (SMSociety). ACM, 2017, 5 pages.

The posters and workshop papers are:

- **Reham Ebada Mohamed** and Sonia Chiasson. [Extended Abstract] Digital aging for increasing privacy in Online Social Networks. Workshop on Exploring Individual Differences in Privacy, The ACM SIGCHI Conference on Human Factors in Computing Systems (CHI), 2018, 4 pages.

- **Reham Ebada Mohamed** and Sonia Chiasson. [Poster] A Study on Aging of Digital Artifacts and its Representation. CLUE Symposium, 2018.

- **Reham Ebada Mohamed** and Sonia Chiasson. [Poster] A Study on Aging of Digital Artifacts and its Representation. SERENE-RISC Fall Workshop, 2017.

- **Reham Ebada Mohamed** and Sonia Chiasson. [Poster] A Study on Aging of Digital Artifacts and its Representation. Canadian Celebration of Women in Computing Conference (CAN-CWiC), 2017.

1.5 Thesis Outline

In Chapter 2, we give background on online reputation, privacy, and existing mechanisms for online reputation management. We also introduce the current research gap that we are addressing. In Chapter 3, we present our preliminary work to understand how online reputation affects OSN users. In Chapter 4, we describe our second study which explored users' understanding and opinions of decay representations and the concept of digital decay in OSNs. In Chapter 5, we present our work on evaluating how decay representations impact users' online reputation in a hiring context. In Chapter 6, we further extend our

evaluation of decay representations using different data collection methods and with Canadian managers. Finally, in Chapter 7, we discuss our findings, summarize our work, and present possible future research directions.

Chapter 2

Background

In this chapter, we provide an overall introduction and background on related topics, such as social media, digital footprints, influence of online reputation on the hiring process, online privacy and online reputation management, existing tools for privacy, and privacy visualizations. We also identify existing research gaps that our research aims to address.

Since the phenomenon of sharing data online is broad and includes various dimensions, some aspects are beyond the scope of our review. Among these dimensions are issues of practical implementation and enforcement of privacy laws. Other issues relate to data collection and behavioural tracking by institutions or apps [23, 77, 114, 148]. While these are important concerns, they are tangential to our current research questions.

2.1 Online Privacy

A dichotomy exists between online users' reported attitudes and their actual behaviour towards privacy, coined as a *privacy paradox* [1, 2, 5, 31, 64, 95, 101, 111, 126]. Online users report willingness to protect their own privacy [2], but studies show that few actions are performed for that purpose [2, 5, 101, 126]. Moreover, even privacy-concerned individuals knowingly disclose information that might be sensitive when they are in specific web contexts, such as online shopping [126], or when expecting a payoff or a reward [2].

The Internet exploits the fact that the privacy paradox exists among users by making salient the desire to divulge while downplaying the desire for privacy [64]. In addition, Coopamootoo and Groß suggest that it may be challenging for users to follow both a *privacy attitude* and a *sharing attitude* simultaneously because the two attitudes stem from two opposing forces or emotions: fear and happiness, respectively [31].

Barth et al. [9] proposed a formal model of privacy and contextual integrity that links protection of personal data to norms in specific contexts. *Contexts* refer to how individuals

act in certain roles within distinctive social domains [9]. The model serves as a conceptual framework endorsing the concept that privacy is not about secrecy, and individuals willingly share personal information if they are assured that specific social norms have not been violated.

2.1.1 Privacy by Design

In 1995, a joint study [135] by the Dutch Data Protection Authority and the Information and Privacy Commissioner of Ontario in Canada explored a new approach to privacy and identity protection, that served as basis for seven *Privacy by Design (PbD)* principles [22], namely:

1. Proactive not reactive; preventative not remedial
2. Privacy as the default setting
3. Privacy embedded into design
4. Full functionality — positive-sum, not zero-sum
5. End-to-end security — full lifecycle protection
6. Visibility and transparency — keep it open
7. Respect for user privacy — keep it user-centric

The PbD principles serve as a framework for proactively embedding privacy during the system engineering process and more broadly within organizational practices. The framework's main goal is to make central the concern for individual privacy by promoting user trust and accountability when handling personal data. PbD principles inform our proposed recommendations in Chapter 7.

2.1.2 Right to be Forgotten

The idea that individuals should be able to move beyond their past artifacts and actions has been most prominently discussed by the EU under principles of the *Right to be forgotten (RTBF)*. The RTBF is part of the EU's new General Data Protection Regulation (GDPR) [108] which came into effect in May 2018. The RTBF provides that search engines must remove links to pages that "appear to be inadequate, irrelevant or no longer

relevant or excessive in the light of the time that had elapsed" when requested by individuals. Concepts of data minimization from the RTBF, however, have been included earlier in the EU Data Protection Directive of 1995 [107].

More recently, the Office of the Privacy Commissioner of Canada (OPC) has included *online reputation and privacy* as one of its strategic priorities between 2015 and 2020 [103]. The OPC has been consulting various sources [103, 104], including the public, to advance in protecting individuals' online reputation. As a result, the OPC has identified legislative approaches and proposed various solutions that balance freedom of expression and protection of individuals' reputation [103]. After finalizing these efforts, the OPC will launch an action plan to put the identified measures into practice [103].

The RTBF has fueled the research community to develop better tools to support users' privacy. In the context of digital artifacts, the RTBF was described as a fundamental need in the literature [76,87,113], which recommends designing for forgetting to enhance online user privacy.

2.1.3 Parallels Between Human and Digital Memories

Reasons for forgetting in human memory can be explained by different principles [118]. One principle, *Trace Decay*, explains how we might fail to trace a memory as time passes, leading to fading or disappearing [26, 118]. Trace Decay is based on *Decay Theory* of immediate memory [118]. Decay Theory suggests that items in immediate memory decay as a function of the passage of time with the cause being some unidentified internal process [118]. On the other hand, research has demonstrated that some memories which have not been rehearsed or remembered are remarkably stable in long-term memory [26].

Although there is agreement on how decay produces forgetting in a general sense, there are many possible interpretations among researchers for what the process of decay actually means [118]. Researchers within cognitive psychology are generally skeptical of decay theory, but a minority of memory researchers are proponents of decay theory [118]. Interestingly, the findings of some researchers who are skeptical towards the theory are hardly explained without recourse to Decay Theory [118]. Nevertheless, research evidence is accumulating in favour of decay [118].

2.2 Social Media

Social Media is an umbrella term that includes various applications [8], such as bulletin boards, computer–supported collaborative work, applications that support exchange of interpersonal information (OSNs), text messaging, or instant messaging.

2.2.1 OSNs and Digital Footprints

OSNs are highly dynamic; they feature user-driven content that is combined with interactivity with other users [125]. OSNs have become popular among individuals for online communication and social interactions. 73% of online adults use at least one OSN site [133], with Facebook as the most popular (2.32 billion monthly active users as of January 2019) [24].

"Digital shadows" or "digital footprints" relate to the traces of information that we produce every day and the concerns about who can access and what can be done to this information [87]. Once information is published online, our digital footprints can be stored indefinitely [30, 87]. In some instances, this may be viewed as beneficial. For example, constituents might believe that having a record of artifacts is important for the accountability and fact checking of politicians. However, most individuals are not subject to such regular scrutiny.

People tend to keep physical artifacts with certain tangible or intangible value [68], and online users also tend to keep and archive their digital artifacts [68, 80]. It is thought that the capabilities of digital technology should be used to eliminate limitations of human memory and to provide a valuable lifelong remembering experience [39, 109]. Therefore, some HCI practices seek to support everyday reminiscing [39, 109], use the web as a personal archive and for information management [80], consider digital inheritance [102], and enable reflection on social relationships [122] or personal past [132].

The literature has shown that maintaining online identity is not an ephemeral act, rather, it is an enduring one [58, 151]. Online users, as individuals in the society, transact in different capacities by managing their online identity. They present themselves in a way that matches current social circumstances [5, 51, 52, 58]. In real life, this maps to *impression management* described by Goffman [46]. Goffman provides a *theatrical* description of

individuals transacting within social interactions. With life as a *stage*, individuals act as *actors* with *performances*, and usually put on a *front* that presents an idealized self depending on the social settings [46]. For example, *"when an individual appears in the presence of others, there will usually be some reason for him to mobilize his activity so that it will convey an impression to others which it is in his interests to convey"-Goffman [46]*. The *front* self is different from the *backstage* self, which represents the real authentic self: *"The 'true' or 'real' attitudes, beliefs, and emotions of the individual can be ascertained only indirectly, through his avowals or through what appears to be involuntary expressive behaviour"-Goffman [46].* Technology and social media complicates individuals' impression management process as users of these platforms have a limited understanding of their audience [86]. Each user in each context has an *imagined audience* [86]. Depending on the situation or interaction, users share different content with different audiences. Meanwhile, social media presents a challenge by flattening different audience (e.g., family, colleagues, close friends) into one (coined as *context collapse*), and OSN privacy settings fail to mitigate this issue [86]. As a way to maintain different imagined audience, users tend to create multiple OSN accounts to express themselves differently on each, or rather perform a self-censorship where they do not post specific content with specific imagined audience [86].

Harper et al. [53] and Hogan [58] explored the concept of identity articulation through time on Facebook. They reflected on how outdated content can resurface, highlighting that social media focuses on "now" even though the associated events may have occurred in the past [53, 58].

Users' privacy protection strategies are limited, and require user engagement. These strategies sometimes include faking data [69] or presenting an unauthentic image of themselves [34, 150], which might compromise their well-being [130]. However, some data is obtained through other users [34, 69, 73, 134] rather than directly, such as tagged photos, posts on another's timeline, or friends' comments. These sources may be beyond a user's control and compromise a user's attempt to manage their online presence.

2.2.2 Social Rewards

There are general societal benefits arising from digital footprints. For example, we can use this information to study human behaviour and social interactions [47, 112]. Sharing content online contributes to individuals' well-being and social interactions [1, 14, 20, 74]. Moreover, directed communication on OSNs can promote social bonding and positive feelings [20] and can facilitate the process of finding and interacting with classmates [1] or maintaining relationships with family and acquaintances [56]. In addition, the use of social media provides individuals with needed social support when they experience negative feelings such as grief [137] or loneliness [74]. Online communication and social media can also positively contribute to adolescent development through increasing self-esteem and providing an outlet for identity experimentation [14, 129].

2.2.3 Negative Impacts

Sharing information on social media is so common that people rarely think deeply about the consequences before posting. However, HCI studies show that many users later regret sharing content online. Furthermore, although users are keen to reveal details about themselves through social media posts [20, 64, 111], their willingness to re-share the same content significantly decreases with time [101]. As noted by Wang et al. [139], Facebook users most frequently regret posting content related to strong sentiment, such as religion, politics, personal issues, offensive content, or personal lies and secrets. Similar regrets were also observed from Twitter users [124]. These regrets were related to the repercussion of the posts in their own lives, which sometimes caused irreversible damage [4, 48]. Several studies have also explored the unintended negative consequences resulting from personal data shared by others, such as through being tagged in others' photos [13].

Repercussions are not limited to personal consequences. OSNs create a central repository for personal information and hence allow for a high level of surveillance, as suggested by Barnes [8]. For example, she describes how an online social exchange between friends can become a way for universities to monitor student behavior [8]. Many incidents and research have demonstrated the potential negative consequences of online sharing on individuals' professional lives. For example, the reputations of health professionals have been impacted as patients search for the digital footprint of their practitioners to evaluate their

professionalism [48]. The negative consequences of online sharing are significant in hiring decisions and this is further discussed in the next section.

2.3 Hiring Decisions in the Digital Age

Users are concerned about information revelation because they fear that a future employer might look at their profiles [150]. Even with privacy management, information can be accessed [28], for example when HR managers access the candidate's private profile [63, 125, 138] (e.g., by sending a friend request to the candidate [125]). Users who take precautions by limiting access to their profiles [147] may be viewed negatively by a recruiter [12]. Users believe that deleting content is hard and that it can be retrieved anyway [147]. This is particularly true when a deleted OSN profile can also appear in an online search outside of that particular platform [28]. We note that although Facebook [1], Twitter [2], and Instagram [3] state that they do not have *ownership* of content, they are granted a transferable license to *use* or *process* content. Processing data (e.g. to create profiling data or to sell to third parties) is arguably more threatening to users' privacy than simply owning data [27]. The three platforms also state that content of deleted/deactivated profiles can be retained in backup copies on their servers for a *limited* amount of time. However, this content may continue to appear when indexed by search engines or when shared by others.

Over the past few years, at least 25% and up to 91% [138] of employers did background checks through OSN data [28]. Nearly half of employers screening OSNs use Facebook data in particular [34]. The number of employers screening candidates through OSNs is especially high in the US [28, 89], where there are fewer legislative or legal limitations when it comes to data protection and right to privacy compared to the EU [28]. In Canada, four jurisdictions have laws that restrict social media checks by employers [142]. Nevertheless, social media content continues to influence individuals' employment status in Canada [146].

In essence, the nature of personal information on OSNs can contribute to discrimination or bias in some cases [63, 69, 125]. It is unclear whether checking OSNs is necessary [33]

[1] https://www.facebook.com/terms.php (last accessed in July 2019).

[2] https://twitter.com/en/tos (last accessed in July 2019).

[3] https://help.instagram.com/478745558852511 (last accessed in July 2019).

or helpful [63] for the employment. Moreover, it may result in unfair decisions as data on OSNs might be inaccurate [33], as explained by online self-representation concerns [28, 36, 116, 130]. For instance, users sometimes fake their online information to present a more favourable or idealized self-image [34, 69] or to restrict strangers from gaining true information [150]. The content that appears in searches may be copies of data deleted by the user that does not reflect them anymore [133]. Further, in many cases, OSN content is missing context that is important for accurate judgment of the individual [63].

Proponents for the use of OSNs data in the hiring process argue that employers should do so to avoid being accused of negligent hiring [127, 133]. But is it appropriate to determine one's professional future based on personal activities? Moreover, is it appropriate to infringe on someone's right to privacy? Ethical and privacy issues [63, 125, 127, 133] have been major reasons against this unfettered access.

2.3.1 Online Content Influences Hiring Decisions

The literature and the media describe several cases in which online content has compromised the professional future of job candidates [28, 73, 87, 119, 125, 138], political candidates [3], or academic students [42, 123]. For example, one candidate was denied a job because recruiters found an old photo of her leading a nonviolent protest in front of a consulate [32]. Another candidate was denied a job as a teacher by university administrators because of a photo of her at a party consuming alcohol [87]. The candidate was wearing a pirate costume and holding a plastic cup that did not in fact show its content [87]. More recently, 20 political candidates [3] faced hardships in their electoral riding because of their online content.

The top reasons in the literature for rejecting a candidate were based on content that shows: (1) concern about publicness of personal content [34], (2) concern about lifestyle [34, 73, 89], personal appearance or portrayal [34], (3) poor communication skills [73, 133], (4) false qualifications [133], (5) inappropriate photos/comments [73, 89, 133], (6) use of alcohol or drugs [73], (7) offence to previous employers or coworkers [73], and (8) low levels of professional appearance [133].

Other studies [34, 73] showed that Facebook data helps predict personality traits. For example, research in management and psychology suggests that personality dimensions

of the Big Five model (extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience) can be predicted through social media content [33,62,69,128,133]. These predictions can be used to inform hiring decisions [33, 69]. In addition to the Big Five dimensions, viewers have used timeline photos to infer if the person is professional, well-grounded, or sociable [34]. Photos of partying or drinking might indicate that the person is not hard working or not suitable for the workplace [34].

2.3.2 Veracity of Online Content

Some recruiters are unconcerned with online reputation and do not give much weight to data available on OSNs because they think that such content might be incomplete, inaccurate, or misleading [73].

Previous HCI work on online self-presentation [36, 116] explored whether a user's curated virtual identity matches their offline identity [130] and others [28] suggest that there is a difference between self-presentation and self-disclosure. Self-presentation is self-data an individual communicates to most others [28]. Self-disclosure is the explicit communication of self-data to which another would not normally have access to [28]. It is suggested that employer probably should have access to self-presentation data but not self-disclosure [28].

Furthermore, interpreting other users' OSN data might not be subjective [125]. For example, if a recruiter is seeking a sales representative, and a desirable trait is that the candidate has a large social network. Looking at a candidate's profile with a high number of friends might favour the candidate over another candidate with fewer friends [125]. This judgment might overlook the fact that the latter candidate is a new user. Hence, the validity of OSN data and the use of such data in critical selection processes is questionable.

2.3.3 Demographics Influence Hiring Decisions

The literature suggests gender differences in self-representation [144] and usage of social media [139]. Female users limit the visibility of their posts more than male users because they are more aware of consequences related to employment (especially pictures with alcohol) and because they are harassed online based on their gender [57].

A study [6] using a fictitious stereotypical profile of a young woman showed that a young woman is not only judged more harshly in terms of content, but also for the degree

of publicness of her online content. Participants expressed concerns related to alcohol pictures when it comes to employment, and that an employer might prefer one that appears more *serious*.

In another employment context, a study [66] showed that female candidates who published content that is considered negative were more harshly viewed by people trained in Human Resources (HR) than when the same content was published by male candidates. This study was performed in the US where laws exist to prohibit the use of gender (and other personal information) of the applicant in the selection process. Such findings suggest that using OSN data results in discriminatory decisions by HR personnel, even if this was unintentional. It further suggests that social constructs related to gender in the workplace still play a role in the selection process.

Managers' demographics may also influence employment and hiring decisions. Differences in attributes such as race [44], gender, or age between managers and employers or potential candidates can influence employment status [43] or hiring decisions [40, 44]. These studies suggested that managers prefer applicants matching their own race/gender/age.

2.4 Tools for Privacy/Forgetting

As discussed in Section 2.2.1, users curate online self-representational data to meet current circumstances [5, 51, 52, 58]. Consequently, usable privacy and HCI literature recognize a need for contextual privacy settings [5, 9, 83, 99]. Previous work showed that Facebook privacy settings did not support users' sharing intentions [83], or their need for reflection [152]. In addition, Novotny and Spiekermann [99] showed that users desire control over their disclosed personal information in OSNs and need to dissociate from obsolete information that represents their past identity.

Confirmed with the privacy paradox [1, 2, 5, 31, 64, 95, 101, 111, 126], users appreciate privacy as a preventive measure, yet they rarely perform active privacy management or online reputation management. This is in part due to a lack of usability, as explained by Yang [147], who recommends that the usability of privacy settings becomes the first priority of OSN providers. Moreover, Woodruff [145] suggests that more usable ORM strategies should focus on prevention mechanisms, minimizing user engagement, and developing feasible solution for everyday reputation problems.

On the other hand, OSN *defaults* promote sharing of information rather than promoting privacy [81,88,143]. These defaults require complex tasks from users to customize, placing a burden on users when modifying their privacy [81, 88, 143]. For example, every timeline post on Facebook (e.g., photo upload, status update, check-in, etc.) is public by default [88], and to override settings, users have to modify the visibility and specify who the viewers are when publishing each post [81]. On Twitter, all tweets can be either public or private where only approved followers have access, and the user cannot chose a subset of users to view specific tweets [92]. It is possible that OSN providers have a business motive to provide poor privacy tools [45]. For instance, it could be to OSN providers' advantage to retain as much material as possible, since it is valuable for targeted advertising [45].

2.4.1 Preventing Regret

One approach to preventing regrettable information disclosure is nudging. For example, Wang et al. [140, 141] introduced three nudges to Facebook users [141]: reminding users about the audience of the post (Figure 2.1), delaying publishing the post (Figure 2.2), and giving feedback regarding content containing strong sentiments (Figure 2.3). Although perceived as beneficial, users started to ignore the nudges within days. Moreover, while users liked the first nudge, they found the second and third nudges intrusive. Another way to prevent potential future regret about disclosed information is to set an expiry date for the published information [101], which is later discussed in this subsection.

2.4.2 Deletion

Deletion [10, 30] permanently erases user online content from respective platforms, which limits users from further access to their content. Ayalon and Toch [5, 101] found that users' willingness to re-share information decreases with time, as it becomes less relevant. In the meantime, the probability that they delete such irrelevant information was low [5, 101] and there was no obvious tendency of users to permanently change their old content, and sometimes they express regret over deletion decisions [152]. Thus, users' reported approaches towards sharing do not align with their actual behaviour, which could be explained by the privacy paradox [5]. However, other reasons also include the desire to keep past posts for reminiscing [5, 10] or reflection [152]. Therefore, mechanisms that permanently delete

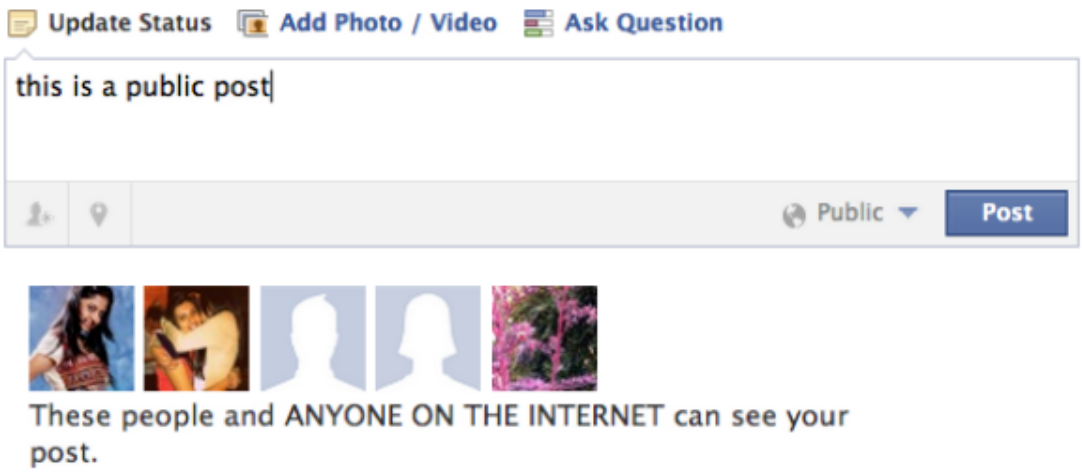


Figure 2.1: Profile picture and audience nudge. Figure reprinted from Wang et al. [141].

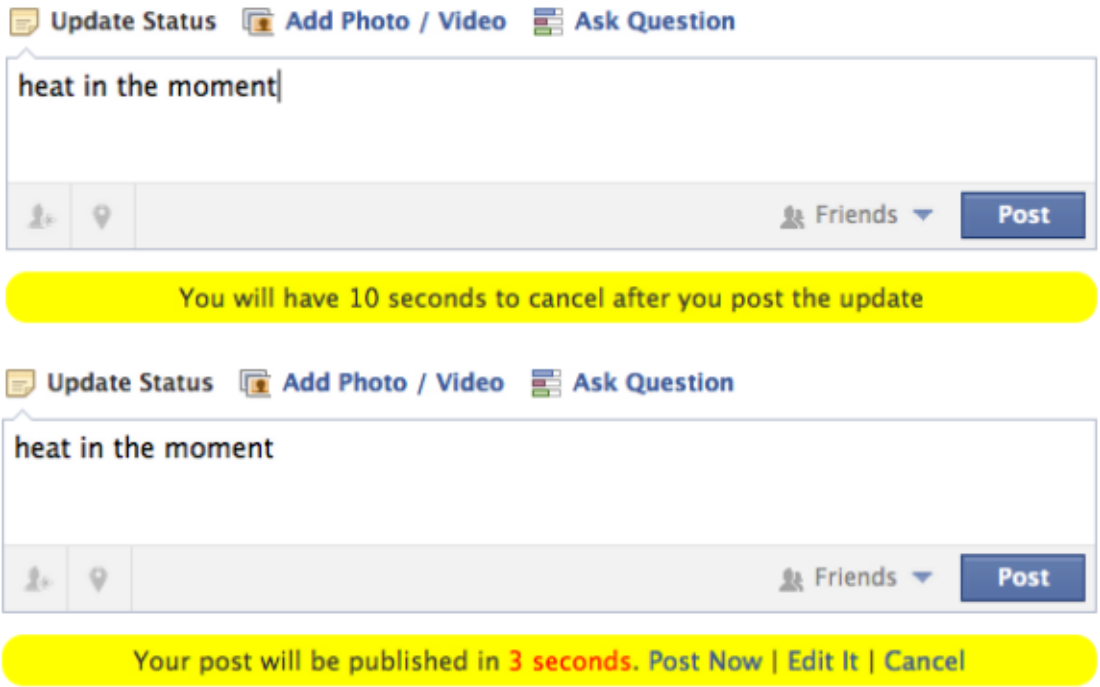


Figure 2.2: Timer nudge before (Top) and after (Bottom) clicking "post". Figure reprinted from Wang et al. [141].

Figure 2.3: Sentiment nudge depending on the sentiment of the post content. Figure reprinted from Wang et al. [141].

content do not appear appropriate for most users as a solution for long-term retrospective privacy [5] or when curating their online-self [152]. These mechanisms include solutions such as the Web 2.0 Suicide Machine [30], or deleting content after a certain amount of time [10]. Permanently deleted content from user view may however continue to appear, for example, when the content had been shared by other users or had been indexed by search engines.

2.4.3 Expiry, Archival, and Decay

We review three forgetting mechanisms to manage online content that could be alternatives to content deletion.

Expiry: The user sets an expiration date for the content at the time of its publication [5, 101]. Upon expiry, the content is removed from view.

Archival: Content is moved from the OSN platform to a different storage location that is accessible by the content owner [80, 151].

Decay: Obsolete content is gradually removed at a particular decay rate that could be automatic or controlled by the user [10].

Based on the identified gap between users' sharing preferences and their willingness to delete, Ayalon and Toch [101] proposed an information expiry feature for Facebook.

They [5, 101] also suggested other mechanisms for ongoing privacy management instead of deletion: archiving, compaction, and blocking.

When considering information expiry, it might be challenging to set expiration defaults to accommodate preferences for sharing information across different time periods [101]. In addition, Bauer et al. [11] cast doubt on the usefulness of content expiration and suggested that extensive archival features would not be appropriate for users. Through two studies about privacy settings using the temporal dimension, Bauer et al. [11] found a gap between users' prediction about how their own privacy preferences would change over time and the actual change in their preferences. They instead suggested designing interfaces that promote reflection on older content [11]. Gulotta et al. [51] suggested that a more subtle mechanism to handle irrelevant content, such as selective archiving rather than extensive archiving, would be more helpful to users because content remains accessible by the content publisher [5]. Previous work [80, 151] has shown that several complexities exist when considering archival features. Users are demotivated to keep an aggregated archive for their digital artifacts and archiving these artifacts depends on the context and platform [80]. More specifically, when it comes to OSN profiles, users might find it difficult to predict the value of their content in the future [5, 11, 80]. Moreover, they think that their OSN content gains meaning from the context and the interactions at the time of sharing [80, 151]. Therefore, Lindley et al. [80] recommend interactions that allow users to revisit older profiles instead of interactions that back up profiles or move them offline. This is supported by Zhao and Lindley's [151] findings which showed that photos on Facebook and Instagram were more likely to be visited by their owner than those saved offline. These results [80, 151] emphasize earlier findings that OSN content is curated as an exhibit and is always in flux [58], suggesting that OSN content is unlikely to have archival value. Ayalon and Toch [101] suggested that the format of the Facebook timeline offers a reasonable starting point for enabling users to review and reflect on old content, and to manage their privacy.

A more concrete and elaborate theoretical proposal of forgetting mechanisms and interfaces was discussed by Barua et al. [10]. They set forth theoretical foundations for the

design of user-controlled forgetting mechanisms in HCI that parallel forms of human forgetting. They discuss the benefits and consequences of implementing five forgetting mechanisms: decay, deletion, compaction, blocking, and archival. For example, they demonstrate that a decay mechanism gradually removing obsolete content would simulate the decay process in human memory [17, 118].

Other studies [84, 85] explored *institutional archiving* of Facebook, however, this type of archiving is tangential to our current research question, which focuses on end-user interactions.

2.5 Privacy Visualizations

Some privacy visualizations are meant to inform users' behaviour towards privacy. For example, Kelley et al. [67] introduced a visual representation of privacy policies inspired from nutrition, warning, and energy labelling. The representation explained how organizations use and share users' personal information, with the goal of helping users make informed privacy decisions [67]. Caine et al. [21] proposed audience visualizations within OSNs that made users' information disclosure decisions more aligned with their information disclosure preferences [21].

2.5.1 Obfuscation

One approach is to fully or partially obfuscate sensitive imagery (photo/video) elements [55, 61, 78, 105] or user attributes [25, 115, 117]. Obfuscating attributes, however, may not be effective against inference[4] attacks [25].

Ragan et al. [115] proposed visual markups to mask personal data in database records when accessed by users. In their scenario, users need access to this data to make informed decisions based on links between data records. The results showed that the quality of decisions was preserved with the use of markups and with only 30% of data content being visible.

[4] As defined by Chen et al. [25]: "The goal of the inference attack is to obtain the value of a user's private (not publicly accessible) attribute of their OSN profile, by analyzing publicly available background information using machine learning techniques."

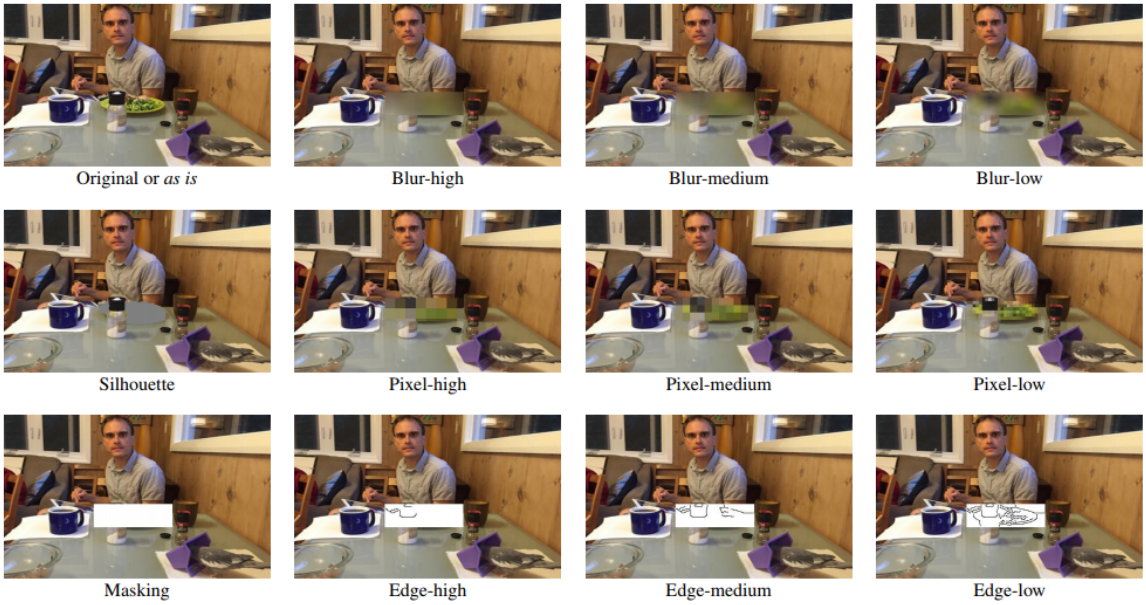


Figure 2.4: Eleven filters applied to food in the original photo. Figure reprinted from Hasan et al. [55].

Padilla et al. [105] provided a comprehensive summary of available imagery obfuscation techniques. Li et al. [78] evaluated the effectiveness of eight obfuscation techniques against human recognition and user experience. Table 2.1 shows the eight tested techniques. Li et al. [78] further showed that some techniques that are commonly used (e.g., pixelation) are neither effective for privacy nor preferred by users. On the other hand, Damanet et al. [35] had previously shown that pixelation could be effective as an obfuscation technique against *human* recognition. For example, increasing both the size of masked area and the pixelation level of deterioration led to lower *human* recognition rates. More recently, Hasan et al. [55] weighed users' perception of both *privacy* and *utility* of five privacy filters that were applied with a different strength (resulting in a total of 11 filters) to obscure an object in photos. *Utility* in this context measured the aesthetics and overall user satisfaction with content viewing experience [55]. Figure 2.4 shows three example filters with three levels of strength obscuring food in the photo. Their findings [55] aligned with previous work, highlighting that increased strength of filters increase perceived privacy, but also decrease utility.

Visual privacy is also widely discussed in the context of video surveillance. In such context, a balance is needed between privacy and intelligibility. If no balance is needed, image

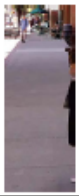
Example	Name & Definition	Related Work	Example	Name & Definition	Related Work
	**Blurring.** Reduces image detail by generating a weighted average of each pixel and its surrounding pixels.	[6, 10, 29, 41, 45]		**Pixelating.** Replaces original small pixels, which are single-colored square display elements that compose the bitmap, with larger pixels.	[24, 39, 40, 45, 80]
	**Silhouette.** Replaces content with a monochrome visual object that mirrors the extracted shape of the original content.	[17, 41, 60, 85]		**Avatar.** Replaces content with a graphical representation that preserves some elements of the underlying content. A human avatar can preserve facial expression and gesture, but hide biometrically unique elements (e.g., face) of identity.	[60, 65, 70]
	**Point-light.** Replaces content with dots that preserves some elements of the underlying content. A human point-light can preserve a person's activity, but hide many biometrically unique elements.	[17]		**Bar.** Replaces content with a monochrome visual object that is the shape of a small, thin rectangle.	[85]
	**Masking.** Replaces content with a monochrome solid box that covers the content to be protected and surrounding image content.	[40, 41, 85]		**Inpainting.** Completely removes content fills in the missing part of the image in a visually consistent manner.	[41, 60, 73, 85]

Table 2.1: Examples of eight obfuscation techniques. Table reprinted from Li et al. [78].

filters such as blurring can be enough to hide identity if gait is not involved as gait and temporal characteristic are difficult to hide when there is familiarity with the target [105]. Padilla et al. [105] indicated that several studies have evaluated obfuscation filters subjectively (through user interviews and questionnaires).

Two related studies [70, 71] included both subjective evaluation concerned with *human* recognition and objective evaluation concerned with *machine* recognition. Both studies evaluated different strengths of blurring, pixelization, and masking in video surveillance, bearing in mind the balance between privacy and intelligibility. They [71] found that an increase in strength of privacy filters leads to an increase in privacy and to a decrease

in intelligibility. They suggest a pixelization block size of 20 can be used in practice. Pixelization appeared slightly better than blurring at preserving privacy, but also negatively affected the intelligibility for larger filter sizes. Overall, masking was the best approach when intelligibility is important such as in video monitoring and surveillance applications.

2.5.2 Decay Representations

One approach to forgetting with digital artifacts visualizes time within the UI. This approach enables users to either manipulate their past content [98] or control how content should be displayed in the future [51]. For example, Gulotta et al. [51] used three systems that incorporated different visualizations, including methods that decay content, to explore how users would like their content to be displayed in the future. Figure 2.5 shows how one system [51] decayed images from their original state to become less meaningful to users.

Temporal Integrity

It is suggested that preserving information's temporal integrity is key in protecting a user's privacy [98]. Temporal integrity could be preserved by indicating the date of online content on the UI. This can be done through manipulating the entire online content [98] or by adding extra time symbols or cues to the UI [98, 100]. An experimental lab study [100] adopted two temporal cues (temporal order and graphical timelines) in a hiring process simulation where reputation profiles of job-seekers were shown to participants acting as employers [100]. Results showed that the graphical timeline helped users more easily disregard obsolete information compared to the temporal order cue.

Novotny's Taxonomy

Novotny [98] looked at dissociating users from obsolete information by having older content gradually decay [98]. This approach can preserve information's temporal contextual integrity [98], which is a key building block of user privacy [16, 94]. Based on results from a focus group, Novotny [98] proposed a catalogue of temporal interface cues to indicate the age of Facebook posts. He classified these cues into *temporal indices* that incorporate time as a property of the posted information and *temporal symbols* that can be used as additional visual cues. The temporal indices manipulate the display properties of the information

Figure 2.5: Digital decay as incorporated in one of Gulotta et al.'s three systems. The process of decay consisted of two stages. In the first stage (middle photo), transparency is increased. In the second stage (right photo), a field of binary 0s and 1s replaces the photo. Figure reprinted from Gulotta et al. [51].

(e.g., through size, motion, decay), while temporal symbols include objects that indicate the time of the post (e.g., adding pictograms) and methods to manipulate the layout (e.g., horizontal or vertical) or typography [98].

A table summarizing Novotny's catalogue is available in Table 2.2, and are briefly explained as follows [98]:

Sedimentation: An analogy borrowed from archaeology, where newer layers of information cover older layers of information on OSNs.

Display salience: Aims to make more recent information stand out.

- Size: Older information is gradually shrinking.
- Motion: Older information moves or flashes in slow motion.

Degrading display quality: Similarly to degradation of physical objects, older information can degrade with time.

- Decay: Image colours are desaturated and vowels in textual information are omitted.
- Greying: Older information is displayed in lighter grey tones.
- Outdated display technology: Older information is displayed using low resolution techniques that imitate older display technologies.

Fashion: An analogy borrowed from fashion trends, where information is represented to reflect its respective time.

- Fashion of content: Information is adapted to the fashion style it dates back to.
- Fashion of UI design: Information is adapted to the look-and-feel of the interfaces it dates back to.

Historic snapshot of the person: The profile picture at the time information was posted is displayed in addition to user's current profile picture.

Symbolic objects: Symbols to be added to the information to indicate its time.

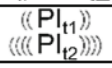
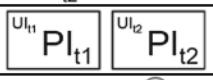
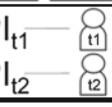
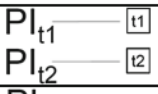
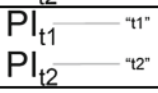
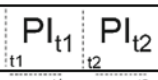
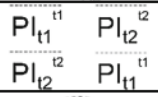
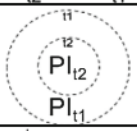
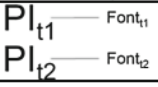
- Time pictograms: Graphical icons or symbols of time-related objects, such as a calendar or clock are displayed next to the information.
- Textual symbols: The user's age at the time information was posted is displayed beneath their name.

Screen space: A way of spatially projecting time into screen areas.

- Horizontal: Older information is placed on the left of the screen.
- Vertical: Older information is placed either on the top or bottom of the screen.
- Concentric: Information grow from the center, placing older information at the perimeter.
- Radial: Information are displayed in radial segments, and older information is placed in an anti clock-wise direction.

Typography: Older information is displayed using classically perceived font styles.

Only one of the temporal indices (*size display salience*) and one temporal symbol (*horizontal screen space*) were partially tested with a small group of users [98]. The prototype [98] gradually degraded the *size* of Facebook posts on a *horizontal* timeline. A snapshot of Novotny' prototype is shown in Figure 2.6.

Temporal sign		Explanation	High time granularity	High amount of PI	Dynamic generation	Generic representation of UI design
Temporal indices						
Sedimentation		Older PI_{t1} is covered by newer layers of PI_{t2}	x	x	x	
Display salience						
	Size	Older PI_{t1} is displayed in smaller size	x	x	x	
	Motion	Newer PI_{t2} moves faster on screen			x	
Degrading display quality						
	Decay	Older PI_{t1} is displayed in decayed state		x	x	
	Greying	Older PI_{t1} is displayed in lighter gray color		x	x	
	Outdated display technology	Older PI_{t1} is displayed using earlier technology		x	x	
Fashion						
	Fashion of content	Fashion of old PI_{t1}'s content is adapted to earlier time		x		
	Fashion of UI design	Older PI_{t1} is displayed using old-fashioned UI design		x	x	
Historic snapshot of the person		Older PI_{t1} is displayed together with old profile picture		x	x	
Temporal symbols						
Symbolic objects						
	Time pictograms	Temporal context is annotated using a graphical symbol	x	x	x	
	Textual symbols	Temporal context is annotated using text or dates	x	x	x	
Screen space		Screen areas containing PI are assigned to time segments				
	Horizontal	Older PI_{t1} is displayed left of newer PI_{t2}	x	x	x	
	Vertical	Older PI_{t1} is displayed on top or bottom	x	x	x	
	Concentric	Newer PI_{t2} is displayed closer to the screen's center	x		x	
	Radial	Radial sections of the screen are assigned clock-wise to newer PI			x	
Typography		Older PI_{t1} is displayed in typefaces perceived as classic		x	x	

PI … SNS profile information, t_i … temporal context at point in time *i*

Table 2.2: Novotny's suggested temporal signs for OSNs. Table reprinted from Novotny [98].

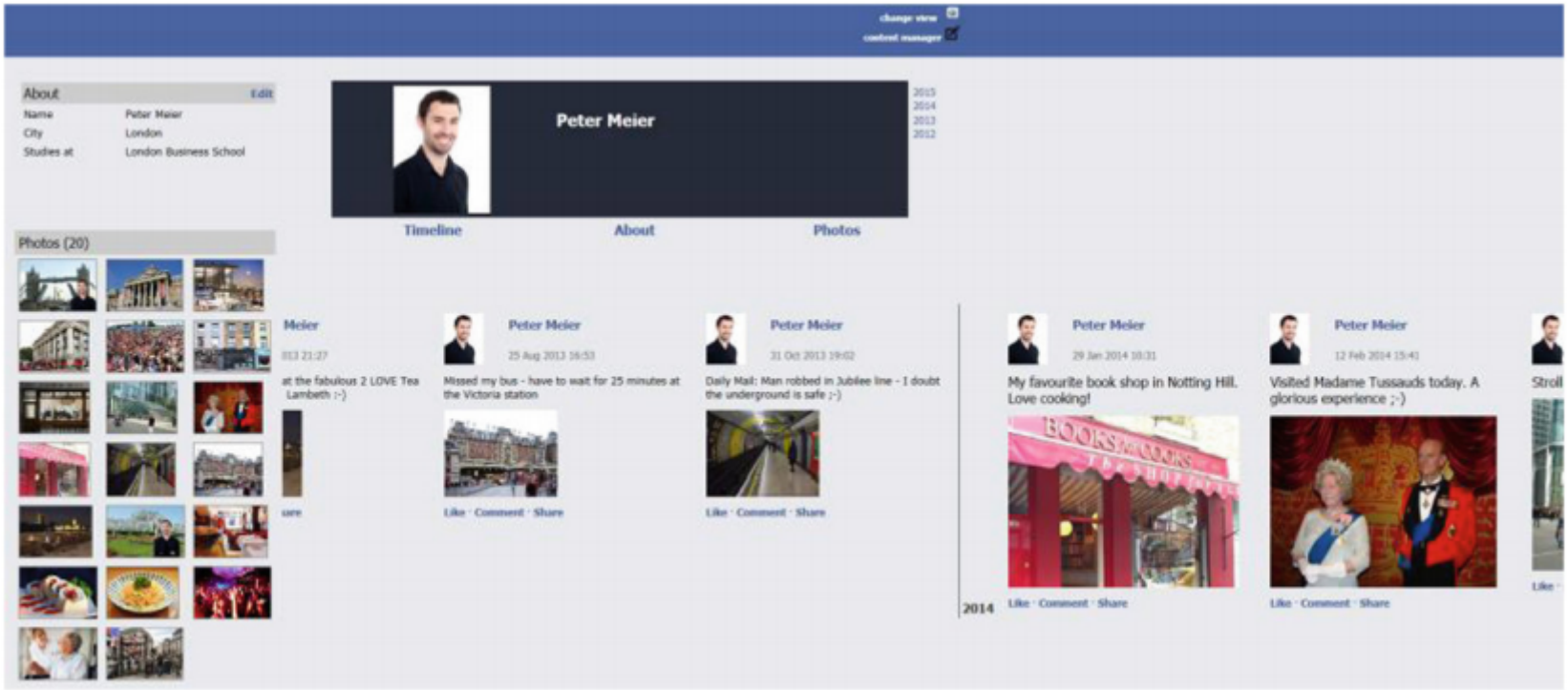


Figure 2.6: A snapshot of Novotny's tested prototype with posts getting smaller with time on a horizontal timeline. Figure reprinted from Novotny [98].

2.6 Existing Gap

We have reviewed current privacy/forgetting mechanisms that have been evaluated, described, or proposed in the literature. Many of these mechanisms have potential benefits but also have obvious concerns as reported by users or as discussed by researchers. We summarize these mechanisms, their potential benefits, and current problems that exist within each in Table 2.3.

The literature has explored both forgetting and reminiscing mechanisms and provided insights on how to better match users' goals, such as providing contextual privacy settings and allowing reflection over older content. However, it is unclear how these mechanisms can co-exist. For example, how can an interface provide an immediate contextual visual cue that can promote privacy whilst presenting a natural non-obtrusive metaphor to users? Another open design and research question in HCI is how the passage of time should be visualized [79]. How do users prefer to depict the passage of time to others, to represent their current personalities, and to show progression in life? What benefits or concerns exist with such mechanisms? And what are the privacy implications relating to these issues? We partially address these research gaps in Chapter 4. Gaps related to privacy implications are also explored in Chapters 5 and 6.

Further, it is unclear how views of online reputation are shaped today. For example, how does users' perception of online reputation impact their assessment of other users in specific contexts, such as for job or political candidates? Moreover, how do managers perceive the online reputation of job candidates in particular? Would decay representations be useful for ORM in such contexts? Chapters 3, 5, and 6 explore these research gaps.

More specifically, each chapter addresses the gaps as follows:

- **Chapter 4:** evaluates decay representations through testing temporal interface cues suggested by Novotny [98]. Decay representations offer an interesting proposal, however, only a few of Novotny's temporal interface cues have been evaluated. In his work, a Facebook prototype visualized the passage of time by gradually decreasing the size of posts, and posts were arranged horizontally on the user's timeline. Although properties of the photo in the post and the caption were manipulated, other

Table 2.3: A summary of reviewed forgetting mechanisms that highlights each mechanism's potential benefit(s) and problem(s).

Forgetting Mechanism	***Use/ Potential Benefit***	***Reported Problem/ Potential Concern/ Current Gap***	***References***
Nudging	-Could prevent regret, especially when posts have a strong sentiment.	-Users found some nudges intrusive.	[140, 141]
Deletion	-Permanently removes user content from online platforms.	-Users expressed regret and wished they could revisit content. -Content may continue to appear (e.g., when shared by others, moved outside the platform, or indexed by search engines).	[5, 10, 30, 101, 152]
Expiry	-Could prevent regret and enables proactive privacy management.	-Users might find it challenging to set appropriate expiry date defaults.	[5, 101]
Archival	-Moves content to a secondary storage. -Content remains accessible by the content publisher.	-Seemed to be complicated with the nature of OSN content. -Users wished to revisit content in its original context at the time of sharing it on the platform itself.	[5, 10, 80, 151]
Obfuscation/ Privacy Filters	-Hides sensitive elements in a photo. -Robust against human recognition, depending on the block size of the applied technique/filter.	-Leaves visible other contextual information that might be revealing.	[35, 55, 61, 70, 71, 78, 105]
Decay Representations	-Look promising in representing the passage of time and preserving a temporal context for the viewer. -Could allow reflection over older content, since content remains accessible by the content owner.	-Have not been sufficiently evaluated.	[98]

contextual cues that might be revealing, such as date of the post [78] were not manipulated. It was also suggested [98] that shrunk posts should still be clickable to ensure readability but it was not clear whether his prototype implemented this feature. However, we think that making the original information available defeats the purpose of degrading them. The prototype was partially evaluated in a study with 14 participants. The horizontal arrangement of posts did not appeal to participants

because it did not match other familiar interfaces which display posts vertically in chronological order. In this thesis, we compared three of Novotny's temporal indices including *size display salience* with 30 participants. We arranged posts *vertically* so that older ones are shifted to the bottom of the timeline. The posts were unclickable to avoid possible retrieval of unmodified content.

- **Chapters 5 and 6:** explore decay representations as a possible tool for ORM by evaluating its acceptability and effectiveness in dissociating users from past online activities. The decay representations have not been previously tested in a social context, such as in an employment situation where the selection process involves managers assessing OSNs profiles of candidates. Furthermore, it is unclear how the use of a decay representation on an OSN profile affects the viewer's impression of the owner. It was suggested that these representations help protect a user's online privacy, but they were neither empirically nor subjectively tested for privacy protection. Previous work [66] suggests that females are more harshly judged. However, these views were reported about a decade ago. It is unclear whether such findings still hold. Furthermore, although it is obvious that OSNs have been informing the hiring decisions for years, it is unclear how the practice is prioritized by current managers. For example, how have priorities shifted in light of recent data protection laws, such as the GDPR? It is also unclear how OSNs shape hiring decision in real-life practices today. Both chapters further explore these gaps.

Chapter 3

The Impact of Digital Footprints in Employment and Political Contexts

In this chapter, we explore how online reputation influence social media users. We gathered opinions of privacy and of how such online content should impact job candidates/current employees and political candidates, respectively. We examined which types of posted content people find unacceptable, whether participants thought recency of the content is relevant, and whether someone's past online behaviour should impact their *professional* or *political* life. Our analysis explored differences between the two contexts (employment and political), and whether demographic characteristics influence users' perspectives towards politicians and/or employees.

Although research has been conducted on related topics [34, 69, 73, 125, 127, 133, 138], we were unable to find any work explicitly comparing whether candidates for jobs or political positions are judged more harshly and how online content might affect their ability to reach their professional goals. The current study partially addresses this gap in the literature and motivated the later phases of this thesis.

The work in this chapter was done in collaboration with Thaís Bardini Idalino. Thaís is a PhD student at the School of Electrical Engineering and Computer Science, University of Ottawa. Both students contributed equally to the study design, data collection, and preliminary analysis. R. Mohamed was more involved in completing and interpreting the statistical analysis. We jointly published it in the ACM proceedings of Social Media & Society (SM&S) 2017 [91].

3.1 Research Questions

Our research questions are:

RQ1: Do participants hold political candidates and job candidates to different standards with respect to online content?

RQ2: Does participants' age/gender/nationality impact views?

3.2 Methodology

We conducted a between-subjects survey with 459 participants to answer our research questions. The study was cleared by our Research Ethics Board.

3.2.1 Questionnaire

We had two parallel versions of the survey: one where the target was a job candidate and one for a political candidate. Both versions were identical except for rephrasing questions to match the context of the survey (job or political). Each had 41 questions according to the following categories.

A. Demographic questions and online behaviour. We collected the participants' gender, age, nationality, education level, occupation, hours spent online, and number of online accounts.

We further asked participants how often they post content that is against their employer's values and beliefs, is embarrassing, is racist, is controversial, is intolerant, is religious, is aggressive, is inappropriate, includes details of their personal life, negatively comments on their current/past employer, and mentions participation in illegal activities. These questions used a 5-point Likert-Scale (never to very frequently).

B. Experience. We asked whether participants had heard of candidates who had their past online activities scrutinized during a job interview or elections, and if they had heard of people being fired or ruled out of elections due to their online content (yes/no questions). We included open ended questions to obtain further details.

C. Perception of online content. We explored participants' perspectives on how different online content should affect candidates. We asked whether the media should dig deep into candidates' online content and whether such content is sufficient grounds for firing an

employee or removing a political candidate. Furthermore, we asked whether they would fire/rule out[1] a candidate who posted such content. Finally, we explored whether recent content should be given more weight than older content. These were 5-point Likert-Scale questions (strongly disagree to strongly agree).

D. Expected consequences. We asked whether online content should affect candidates' professional life or reputation even if it was posted during their non-work hours or before becoming an adult, and whether it should, or does, hinder their professional goals. These questions used 5-point Likert-Scales (strongly disagree to strongly agree).

The list of questions is available in Appendix A. Each question had a "prefer not to answer" option. To check if participants were answering carefully, we added verification questions. We removed surveys with incorrect responses to these questions before analysis.

3.2.2 Participants

Data was collected between November 2015 and April 2016. We launched the surveys through Crowdflower[2] (now renamed Figure Eight), a crowd-sourcing website. Crowdflower workers are ranked according to their history of completing tasks. We excluded participants who have not reached level 3 rank due to low past performance. We opened the survey only to workers from Canada, US, and UK since the political process in these countries is somewhat similar and all had sufficient numbers of English-speaking citizens. To further ensure that workers were paying attention, we required them to spend at least three minutes completing the survey. Crowdflower participants received $0.50US for completing the survey.

At the time of data collection, federal elections had just taken place in Canada. We also launched both surveys through Google Forms and publicized the URL through social media to reach additional Canadian participants.

We collected 847 completed surveys (107 from Google forms and 740 from Crowdflower). After validation of responses and removal of surveys with missing data, we had

[1] In the questionnaire, we use the term "rule out" when considering political candidates. Our intention was that the term would mean *removing* a political candidate from the elections so that they *are unable to compete for their riding*, which was intended as a parallel to "firing" an employee so that they *lose their job position*. In retrospect, this wording was ambiguous and could have been phrased more clearly.

[2] `https://www.figure-eight.com/`

Table 3.1: Participants' gender, age, and nationalities.

	Category	*Job*	*Political*
Gender	Male	150	97
	Female	141	71
Age	Minimum	17	17
	Maximum	74	80
	Mean	37	32
	Std Dev.	14	13
Nationality	Canada	79	69
	US	146	72
	UK	66	27

459 valid surveys for our analysis (291 job surveys and 168 political surveys). Unfortunately, Crowdflower cannot block a worker from answering multiple surveys from the same research team. Thus, we kept only the first survey by any worker which explains the high drop rate between collected responses and the valid ones for analysis. As a result, we had uneven numbers per condition. Table 3.1 lists participant demographics. Most participants had at least a high-school diploma and were moderate to active online users.

3.3 Analysis and Results

Before turning to our research questions, we address a methodological issue that impacted how we analyzed our data. Our surveys had issues in the wording of questions. Specifically, some questions considered *job candidates* while the remaining questions considered *employees* and whether they should be fired. We acknowledge that these are two distinct situations. We only recognized this discrepancy after we had completed data collection and the most reasonable option was to analyze these questions separately by dividing questions into Group Y considering job candidates and Group Z considering employees. Table 3.2 lists the questions in each group from the job survey.

Table 3.2: List of questions that considered *job candidates* (Group Y) and those that considered *employees* (Group Z) in the job survey.

	Questions
Group Y: Job Candidates	**C1:** Employers should dig deep into past online social activities as a mean of evaluating job candidates. **C2:** If I have a job interview, I feel comfortable with employers analyzing my past online activities. **C3:** I understand why employers analyze candidates' past online activities. **C19:** As an employer, I would not interview a job applicant who had inappropriate content online if it was published more than 10 years ago. **C20:** As an employer, I would not interview a job applicant who had inappropriate content if it was published less than 5 years ago. **C21:** As an employer, I would not interview a job applicant who had inappropriate content online if it was published very recently (within the last year). **D1:** I think the online social network content of job candidates does not affect their reputation. **D3:** Should online content affect someone's ability to get a job? **D4:** Does online content affect someone's ability to get a job?
Group Z: Employees	**C4:** Employers should fire employees who post pictures in embarrassing situations. **C5:** Employees whose online activities could damage the company's reputation should be fired. **C6:** Employees whose online activities are against the company's values and beliefs should be fired. **C7:** Employees who express racist comments should be fired. **C8:** Employees who express views against specific groups of people should be fired. **C9:** Employees who express controversial views should be fired. **C10:** Employees who comment on controversial subjects should be fired. **C11:** Employees who express intolerant views should be fired. **C12:** Employees who express religious views should be fired. **C13:** Employees who post aggressive comments should be fired. **C14:** Employees who post inappropriate content should be fired. **C15:** Employees who share details of their personal life should be fired. **C16:** Employees who comment negatively about their current employer should be fired. **C17:** Employees who comment negatively about past employers should be fired. **C18:** Employees who post about participating in illegal activities should be fired. **C22:** Which past online activities really matter in terms of reputation? **D2:** I think people should refrain from publishing content on their online social network that might affect their future image. **D5:** Should details posted by someone during non-work hours affect their professional life? **D6:** Should people be able to have a public online personal life that is separate from their professional life? **D7:** Should content posted before someone legally became an adult be considered by employers? **D8:** Does content posted before someone legally became an adult influence employer decisions?

We believe that the wording of their parallel questions in the political survey lined up and could be compared within their respective group. To illustrate, the questions that considered *interviewing job candidates* (Group Y) in the job survey explored *electing political candidates* in the political survey. The questions that considered *firing employees* (Group Z) in the job survey looked at *ruling out* political candidates in the political survey.

In the following subsections, we report the results of statistical analysis. Some Likert-scale questions were negatively worded. For analysis, we ensured that a score of 1 was always assigned to the most negative response (i.e., were less tolerant of the candidate/employee's behaviour) and 5 to the most positive response (i.e., were more accepting of the candidate/employee's behaviour).

In total, we had nine questions that considered a *job candidate* rather than an *employee*, as listed in Table 3.2. We treated these nine questions separately in our analysis. This is discussed further in Section 3.5.

We used Mann-Whitney U tests to compare components (described below) between the two surveys. To examine if age, gender, or nationality influenced responses, we used Ordinal Regression, Mann-Whitney (MW), and Kruskal-Wallis (KW) respectively. Tests assume a significance level of $p < 0.05$ unless otherwise noted. Given that our data was non-parametric, we chose the non-parametric MW and KW tests.

3.3.1 Principal Component Analysis

As in earlier work [75], we conducted Principal Component Analysis (PCA) to identify the most significant variables (questions) and group them into *components* based on similar response patterns. Previous work similarly used parametric tests with ordinal data [66, 96]. Norman [96] argued that parametric tests are robust in regards with violations posed by non-parametric data. We recognize that opinions diverge around the use of parametric tests with non-parametric data. However, Principal Component Analysis seemed to be our best option. We conducted this analysis on each survey separately, considering 30 questions from Sections C and D of the survey to reduce the questions to a smaller number of variables for subsequent analysis. We ran PCA on two groups of questions separately: 9 questions that considered job candidates (Group Y) and 21 questions that considered employees (Group Z), and their corresponding political equivalents.

Table 3.3: Rotated Component Matrix for Group Y questions in the *job* and the *political* surveys. Rotation was based on *varimax* method. Coloured rows highlight questions that belonged to components common across both surveys.

Question	***Job***		***Political***			***Assigned Component***
	1	***2***	***1***	***2***	***3***	
C1: digging into past content as a mean of evaluation	0.550		0.845			
C2: being comfortable with analysing past online content	0.698		0.846			
C3: understanding why online activities are analysed	0.802	0.725				
C19: more than 10 years		0.880		0.860		
C20: less than 5 years		0.880		0.911		
C21: within the last year	0.555	0.520		0.706		
D1: OSN content does not affect reputation	0.608				0.751	
D4: online content does affect ability to get a job	0.629				0.853	
D3: online content should affect ability to get a job	0.717					

As expected, the resulting Rotating Component Matrix (showing how variables are grouped into components) was different in each survey. We ignored correlations of less than 0.5. Reliability Analysis tests on the components were used to confirm how closely each set of questions are related as a group. Previous literature recommends using discretion when deciding on acceptable alpha value and that it is dependent on the research context [110]. In general, 0.6 is an acceptable value and an alpha below 0.6 (or greater than 0.95) is undesirable [82, 110].

Following the reliability test, we computed a component score by calculating the average value of Likert scores for the included questions, giving each component a score between 1 (least positive) and 5 (most positive). These component scores were used for the further analysis.

We describe below the PCA results for Group Y, followed by its results for Group Z. The descriptions of components (CM) below are phrased in terms of the job survey, with modifications for the political survey in parentheses.

Results for Group Y:

We extracted groupings common across both surveys, resulting in 8 questions grouped into three components. Cronbach's alpha for all components in both surveys was above 0.6. Table 3.3 shows the Rotated Component Matrix for the Group Y questions of *job* and the *political* survey which illustrates how questions loaded into components.

Table 3.4: Rotated Component Matrix for Group Z questions in the *job* and the *political* surveys. Rotation was based on *Varimax* method. Coloured rows highlight questions that belonged to components common across both surveys, with excluded components highlighted in shades of gray.

Question	*Job*				*Political*					*Assigned Component*
	1	*2*	*3*	*4*	*1*	*2*	*3*	*4*	*5*	
C5: could damage reputation		0.732			0.824					
C6: against beliefs		0.656			0.782					
C18: illegal activities		0.676			0.700					
C16: comments negatively about current employer		0.739				0.725				Y-CM2
C17: comments negatively about past employer	0.585	0.518				0.868				
C9: expresses controversial comments	0.767					0.726				
C10: comments on controversial subjects	0.844					0.815				
C15: shares details of personal life	0.723					0.729				
C7: racist comments			0.852		0.640		0.567			
C8: against groups			0.815		0.502		0.633			
C11: intolerant views			0.727				0.799			
C4: pictures in embarrassing situations					0.510					
C12: religious views	0.633									
C13: aggressive comments	0.531									
C14: inappropriate content					0.721					
D2: should refrain from publishing content that might affect future image		0.509								
D5: non-work hours should affect professional life				0.743				0.774		EXC1
D6: should have online life				0.666				0.786		
D7: content posted before legally adult should influence				0.721					0.737	EXC2
D8: content posted before legally adult does influence				0.620					0.767	

Y-CM1: Content Investigation. Questions C1, C2, and C3 explored whether respondents were comfortable with having past online activities of candidates investigated by employers (media).

Y-CM2: Time Span. Questions C19, C20, and C21 asked if content posted (i) 10 years, (ii) 5 years, and (iii) less than one year ago would impact whether the respondent interviewed (voted for) the candidate.

Y-CM3: Consequences. Questions D1 and D4 explored whether online content affects candidates' reputation or ability to get a job (be elected).

Results for Group Z:

We extracted groupings common across both surveys, resulting in 15 questions grouped into six components. We excluded two of the six components after running a reliability analysis. Cronbach's alpha for four components in both surveys was above 0.7. The four components grouped a total of 11 questions. We excluded the remaining two components because their alphas in the political survey were below 0.6. Table 3.4 shows the Rotated Component Matrix for the Group Z questions of the *job* and the *political* survey.

Z-CM1: Reputation of the Company/Country. Questions C5, C6, and C18 asked if candidates should be fired (ruled out of elections) when their online activities (i) could damage the company's (country's) reputation, (ii) went against the company's (country's) values, or (iii) were illegal.

Z-CM2: Offending Employers/Parties. Questions C16 and C17 asked if candidates should be fired (ruled out of elections) when they comment negatively on their (i) current or (ii) past employer (political party's policies).

Z-CM3: Controversial Content. Questions C9, C10, and C15 asked if candidates should be fired (ruled out of elections) when they (i) express controversial views, (ii) comment on controversial topics, or (iii) share personal details.

Z-CM4: Discriminatory Content. Questions C7, C8, and C11 explored if candidates should be fired (ruled out of elections) when they express (i) racist comments, (ii) views against specific groups, or (iii) intolerant views.

The PCA result was generally as expected: the questions included in each component were thematically related and could plausibly be grouped. One unexpected grouping was Z-CM3, where questions relating to personal details and controversial data were grouped under one component. We believe this may be because some participants interpreted the term 'personal details' as 'intimate' or otherwise more 'controversial' than we had initially intended.

3.3.2 RQ1 Analysis and Results

RQ1: *Do participants hold political candidates and job candidates to different standards with respect to online content?*

We address our first research question, RQ1, by comparing the responses to the job and the political surveys using Mann-Whitney (MW) tests on the resulting components from Groups Y and Z, with Bonferroni correction. Given that our data was non-parametric, we use the non-parametric MW tests to answer our first research question. Table 3.5 summarizes descriptive statistics for *Group Y components* and *Group Z components* respectively. In this, and subsequent tables, lower means are in bold-red and pairs where significant differences were found have grey backgrounds.

Table 3.5: Descriptive statistics for Group Y and Group Z components in each survey. Grey cells indicate a significant difference and means in bold-red represent the less tolerant view.

Question	*Job*			*Political*		
	Median	Mean	SD	Median	Mean	SD
Y-CM1: Content investigation	2.7	2.8	(0.9)	2.7	**2.5**	(0.7)
Y-CM2: Time Span	3.3	3.3	(0.9)	3.0	**2.8**	(0.9)
Y-CM3: Consequences	2.5	2.5	(0.8)	2.0	**2.2**	(0.8)
Z-CM1: Reputation	2.7	2.7	(0.8)	2.7	**2.6**	(0.9)
Z-CM2: Offending employers/parties	3.5	**3.3**	(0.8)	4.0	3.7	(0.8)
Z-CM3: Controversial content	4.0	4.0	(0.7)	3.7	**3.7**	(0.8)
Z-CM4: Discriminatory content	3.0	2.8	(0.9)	2.5	**2.6**	(1.0)

We found a significant difference in all three components from Group Y where respondents were less tolerant of political candidates. In addition, we found a significant difference in three out of four components from Group Z. In two components (Z-CM3, Z-CM4), respondents were less tolerant of political candidates than job candidates and vice versa for Z-CM3. We report details of these results below.

MW tests with significant results for Group Y:

Y-CM1: Content Investigation. Political survey respondents were more inclined to investigate the online content of candidates than those assessing job candidates ($p = 0.003$).
Y-CM2: Time Span. Political survey respondents were less tolerant of content posted in any time span (less than one year, 5 years, or 10 years ago). For instance, they were less likely to elect a candidate with inappropriate online content published up to 10 years ago than those evaluating job candidates ($p = 0.000$).
Y-CM3: Consequences. Political survey respondents believed more strongly that the online content of candidates does affect their reputation and ability to be elected compared to those assessing job candidates ($p = 0.000$).

MW test with significant results for Group Z:

Z-CM2: Offending Employers/Parties. Job survey respondents believed more strongly in firing employees who post online content that is offensive to their employer compared to political candidates offending their political party ($p = 0.000$).

Z-CM3: Controversial Content. Political survey respondents were more inclined to rule out political candidates who post any controversial content, than job survey respondents assessing job candidates ($p = 0.000$).

Z-CM4: Discriminatory Content. Political survey respondents were more strongly in favor of ruling out political candidates who post online content that is discriminatory than those assessing job candidates ($p = 0.002$).

We could also have compared within components to explore whether participants felt more strongly about some components than others, or to examine which components contributed more to firing/ruling out employees or political candidates in their assessment. However, the questionnaire was not explicitly designed to compare within components; we ensured parallel wording between the surveys to only explore differences between the two scenarios, not that individual components were comparable to each other within scenarios.

3.3.3 RQ2 Analysis and Results

RQ2: *Does participants' age/gender/nationality impact views?*

To address research question RQ2, we investigated the effects of respondents' age, gender, and nationality on responses. We use non-parametric test when exploring gender and nationality effects since our data was non-parametric.

Age. We used Ordinal Regression to explore whether age affected the seven components from Groups Y and Z. Results showed no significance on all components within the *job* survey except *A-CM1: Content Investigation* ($p = 0.015$). Using odds ratios (OR) based on the beta values (estimates), we found that older participants were less tolerant than younger participants of *job candidates*, believing that their online content should be considered by employers ($OR = 1$).

There was a significant effect of age on four components in the *political* survey; *A-CM1: Content Investigation* ($p = 0.025$); *B-CM1: Reputation* ($p = 0.001$), *B-CM2: Offending employers/parties* ($p = 0.003$), *B-CM3: Controversial Content* ($p = 0.000$). Using odds ratios (OR) based on the beta values (estimates), we found that older participants were less tolerant than younger participants of *political candidates*, believing that their online content should be considered during elections (OR=1). They were also less tolerant of political candidates posting content that might *offend* (OR=1) their political party, or that might be

Table 3.6: Descriptive statistics for components in each survey divided by *gender*, with significant results highlighted. 1 = least tolerant, 5 = most tolerant.

CM#	***Job***						***Political***					
	Male			Female			Male			Female		
	Md	μ	SD	Md	μ	SD	Md	μ	SD	Md	μ	SD
Y-CM1	2.7	2.8	(0.9)	2.7	2.8	(0.9)	2.7	2.5	(0.6)	2.3	2.5	(0.8)
Y-CM2	3.3	3.4	(0.8)	3.3	3.4	(0.8)	3.0	2.9	(0.7)	3.0	2.9	(0.9)
Y-CM3	2.5	2.6	(0.8)	2.0	2.3	(0.7)	2.0	2.2	(0.8)	2.0	2.2	(0.8)
Z-CM1	2.7	2.8	(0.8)	2.7	**2.6**	(0.7)	2.7	2.7	(0.8)	2.3	**2.4**	(0.8)
Z-CM2	3.5	3.4	(0.8)	3.5	3.2	(0.7)	3.5	3.7	(0.8)	4.0	3.8	(0.8)
Z-CM3	4.0	3.9	(0.7)	4.0	4.0	(0.7)	3.7	**3.6**	(0.8)	4.0	3.9	(0.7)
Z-CM4	3.0	2.8	(0.9)	3.0	2.8	(0.9)	2.7	2.6	(1.0)	2.3	2.4	(1.0)

considered *controversial* (OR=1). On the other hand, they were more tolerant of political candidates posting content that might *threaten the image* of the country (OR=0.96).

Gender. We used independent samples Mann-Whitney tests to examine if gender affected the seven components. Descriptive statistics per gender are shown in Table 3.6.

In the job survey, we found no significant effect of gender, except for *Y-CM3: Consequences* ($p = 0.001$) and *Z-CM1: Reputation* ($p = 0.026$). Male participants were less likely to believe that online content affects *job candidates* than female participants. Female participants were less tolerant of *employees* who post content that compromises the reputation of the company than male participants.

In the political survey, there was a significant effect of gender on *Z-CM1: Reputation* ($p = 0.011$) and *Z-CM3: Controversial Content* ($p = 0.004$). Female participants were less tolerant of political candidates who compromise the country's image. However, male participants were less tolerant of political candidates who post controversial content.

Nationality. We used independent samples Kruskal-Wallis tests to examine if nationality affected the seven components. Descriptive statistics per country are shown in Table 3.7.

A significant effect of nationality was found on both surveys for two components: *Z-CM3: Controversial Content* (Job: $p = 0.003$, Political: $p = 0.000$), and *Z-CM4: Discriminatory Content* (Job: $p = 0.023$, Political: $p = 0.007$). Additionally, a significant effect was found on *Z-CM2: Offending Employers/Parties* on the political survey ($p = 0.000$).

We followed up with Bonferroni-corrected pairwise comparisons (significance level

Table 3.7: Descriptive statistics for components in each survey divided by *nationality*, with significant results highlighted. 1 = least tolerant, 5 = most tolerant.

CM#	***Job***									***Political***								
	CA			UK			US			CA			UK			US		
	Md	μ	SD	Md	μ	SD	Md	μ	SD	Md	μ	SD	Md	μ	SD	Md	μ	SD
Y-CM1	3.0	2.9	(0.8)	2.7	2.8	(0.9)	2.7	2.8	(0.9)	2.7	2.6	(0.6)	2.3	2.4	(0.5)	2.3	2.4	(0.8)
Y-CM2	3.3	3.4	(0.6)	3.3	3.6	(0.8)	3.3	3.4	(0.8)	3.0	2.9	(0.6)	3.0	3.1	(0.9)	3.0	2.8	(0.8)
Y-CM3	2.5	2.4	(0.7)	2.5	2.5	(0.8)	2.5	2.5	(0.8)	2.0	2.0	(0.6)	2.5	2.4	(0.7)	2.0	2.3	(1.0)
Z-CM1	2.7	2.8	(0.6)	2.7	2.6	(0.9)	2.7	2.6	(0.8)	2.7	2.6	(0.7)	2.7	2.8	(1.0)	2.3	2.4	(0.9)
Z-CM2	3.5	3.3	(0.6)	3.5	3.3	(1.0)	3.5	3.3	(0.8)	3.5	**3.4**	(0.6)	4.0	4.2	(0.8)	4.0	3.8	(0.8)
Z-CM3	3.7	**3.8**	(0.6)	4.0	4.1	(0.6)	4.3	4.0	(0.8)	3.3	**3.5**	(0.5)	4.0	4.2	(0.7)	4.0	3.8	(0.9)
Z-CM4	3.0	2.9	(0.8)	2.3	**2.6**	(0.9)	3.0	2.9	(0.9)	2.7	2.6	(0.8)	3.0	3.0	(1.0)	2.0	**2.3**	(1.0)

Table 3.8: Effect of nationality. Asymp. Sig. values as reported from the KW pairwise comparison; Values with Bonferroni-corrected significant differences are highlighted in gray.

Component	***CA-US***	***CA-UK***	***UK-US***
Job, Z-CM3: Controversial Content	0.007	0.012	1.000
Job, Z-CM4: Discriminatory Content	1.000	0.072	0.026
Pol, Z-CM2: Offending Employers/Parties	0.002	0.000	0.139
Pol, Z-CM3: Controversial Content	0.007	0.000	0.214
Pol, Z-CM4: Discriminatory Content	0.125	0.419	0.008

of $p < 0.05/3$, i.e., $p < 0.017$), as shown in Table 3.8. Canadians were less tolerant of employees and political candidates who post *controversial* content. Further, they were less tolerant of political candidates who post content *offensive* to their political party. Finally, responses from the US were less tolerant of political candidates who post *discriminatory* content compared to responses from the UK. This is especially interesting given the current political climate in the US. Although Z-CM4 showed a significant effect for the job survey, the subsequent pairwise comparisons were not significant after the Bonferroni correction.

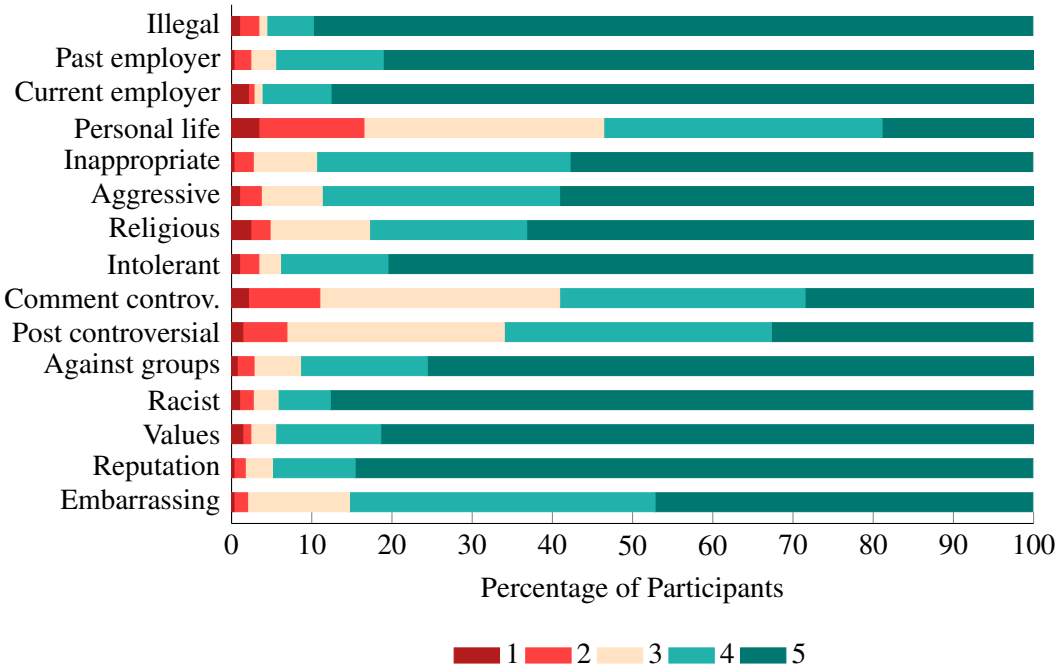


Figure 3.1: Job Survey: Frequency of posting specific categories of online content (1 = very frequently, 5 = never).

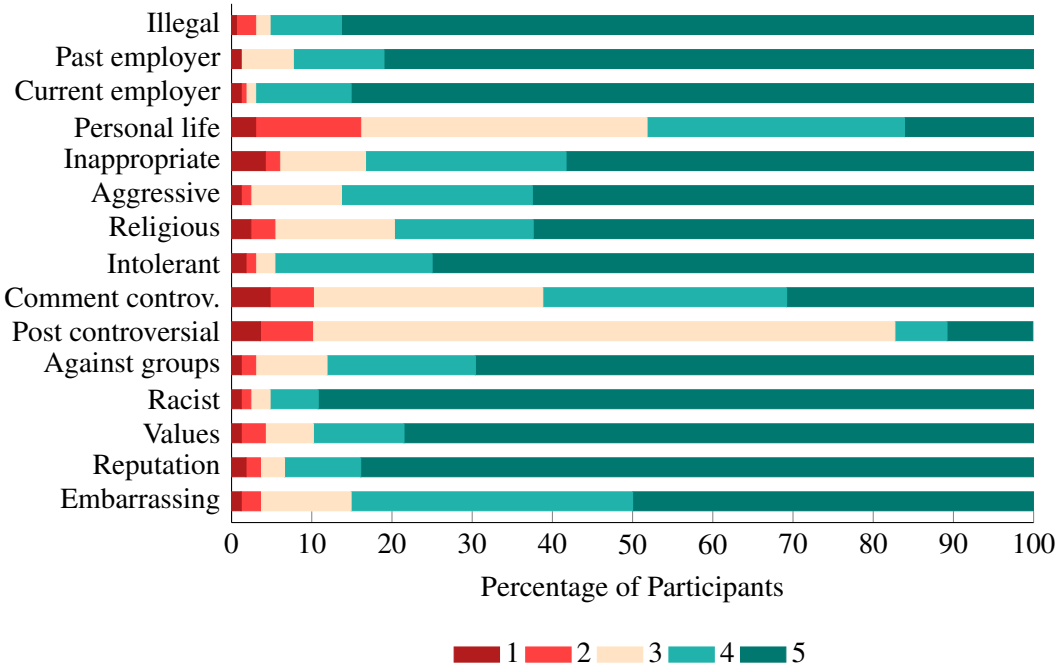


Figure 3.2: Political Survey: Frequency of posting specific categories of online content (1 = very frequently, 5 = never).

3.3.4 Participants' Own Online Behaviour

To put participants' responses into context, we also asked them about their own habits[3] in posting online content (15 questions from Section A of the questionnaire). Figures 3.1 and 3.2 summarize their responses. As shown, participants in both groups reported similar posting habits. We ran Mann-Whitney U Tests to confirm that these habits were similar. We found no significant difference across all 15 questions except for one that considered posting *controversial* content ($p = 0.000$). Participants from the job survey were more likely to post controversial content ($\mu = 3.9, SD = 1.0$) compared to participants from the political survey ($\mu = 3.1, SD = 0.8$).

3.4 Discussion

Addressing our two research questions, we found that participants were considerably less tolerant towards political candidates, holding them to a higher standard with respect to their digital footprint than current employees or even job candidates (RQ1). For RQ2, we found that older respondents were less tolerant than younger participants of political candidates' online activities on three of seven components and more tolerant on a fourth component. Gender also impacted respondents' opinion of political candidates' online activities, but no clear pattern emerged. Nationality impacted responses on three of seven components, with Canadian respondents being less tolerant. There were far fewer differences in the job survey.

The results seem plausible considering two elements that may influence respondents' opinions. First, politicians are meant to represent their constituents to the outside world and act as an advocate for their citizens. It is sensible that citizens want to elect politicians with a history of responsible and mature behaviour, and who have limited potential for causing scandal. Moreover, as public figures, the line between politicians' professional and personal lives is blurred. We note that the surveys were conducted just after the 2015 Canadian election where the media uncovered questionable material on a number of local candidates prior to the election which resulted in candidates withdrawing or losing their

[3]We only collected responses from participants who have reached level 3 rank which indicates a high past performance. However, it is plausible that a social desirability bias [49] affected their responses when reporting their own habits.

respective electoral ridings. The 2016 US presidential election had not yet taken place. Hence, the timing of the survey may have influenced responses. It would be interesting to repeat the survey given this new political climate as we anticipate far more polarizing results, particularly divided along political party lines. This survey provides an interesting snapshot immediately preceding major political events and can serve as a benchmark against which new data can be compared.

Secondly, most respondents are unlikely to ever become a political candidate, therefore, it may be easier to hold such candidates to higher standards. Respondents do not need to consider how their own online footprint might impact their chances at elected office. On the other hand, respondents may be more sympathetic towards job candidates or current employees since they have been in a similar position; respondents may have questionable social media content and recognize that they would like a 'second chance' if it might impact their job prospects. We also note that there may be a considerably higher tolerance threshold for employees compared to job candidates, since firing someone is likely considered more drastic that overlooking someone in a hiring process. Our survey does not provide data for this comparison, however.

We believe that this survey raises interesting questions about social media, privacy, and digital footprints. As more of our lives are digitized, we, as a society, need to consider the implications. It is unlikely that anyone considered a 'digital native' will reach middle-age without having some questionable content in their digital footprint. Do we become more tolerant of such content, should questionable content disqualify someone from attaining later professional goals, do we work towards technical solutions that give users more control to erase their footprint and control their privacy (e.g., the 'right to be forgotten'), or do we put in place mechanisms that automatically delete data after a given time? Each approach has its own implications and impacts. Interestingly, this study highlights situations where social media content may have short-term 'good' consequences, but can lead to 'bad' long-term outcomes for users. On the other hand, social media makes it easier for others to assess a candidate's character, assuming that posted content is legitimate.

3.5 Limitations

Although generally accepted within the usable security and HCI communities, crowdsourcing data may have biases. Furthermore, we recruited some of our Canadian data through different methods. We believe, however, that the general trends observed are reasonable and likely reflect the wider population. Additionally, the recruitment methods potentially skewed the results towards users who are familiar/comfortable with online platforms. Further study could explore whether this generalizes to other populations. Since we asked about two different contexts, it was necessary to reword some questions. We tried to make them as parallel as possible, but it is possible that the wording changes impacted some responses. Moreover, because some questions explored effect on job candidates and other questions considered firing current employees, this discrepancy might have skewed responses. We tried to mitigate this issue by treating the two sets of questions separately. It would be worth exploring responses with another survey focusing on only one aspect.

3.6 Conclusion

This chapter presented a survey investigating how social media activities and digital footprints affect job candidates/current employees and political candidates. We found significant differences in how these groups of individuals were evaluated by respondents. Respondents' demographics had limited impact on the results. Results reflect incidents reported in the media where some politicians had their career compromised by their digital footprint [3, 29], and research on the impact of online content on professional careers [87]. This study provides insights into how online activities affect reputation and it was the first to directly compare evaluation of job and political candidates based on their online behaviour. It has also raised interesting questions about the associated societal impact of our growing digital footprints.

Chapter 4

Decay Representations for Increasing Online Privacy

Exploring potential technical mechanisms to address the long term negative consequences of social media content seen in Chapter 3, we present our work on *decay representations* (introduced in Section 2.5.2). *Decay representations* gradually decay content so that it becomes less accessible to audiences. We believe this is a gentler approach to simulate the idea of archiving and acts as a subtle mechanism to handle digital artifacts (discussed in Section 2.4.3). It also provides an immediate contextual cue to the viewer about the age of posted content.

We extend Novotny's [98] study by comparing three different representations (pixelating, fading, and shrinking) on three different OSN platforms (Facebook, Instagram, and Twitter). We choose three different OSN platforms instead of one to see if our findings are applicable across platforms. We also choose three distinct representations that degrade content differently and fall under two of Novotny's [98] suggested temporal indices: *display salience* and *degrading display quality*. We explore participants' attitudes and concerns about how *decay representations* relate to protecting their online privacy. Our study partially answers some of the open research questions regarding visualizing time in OSNs.

The work in this chapter was published in the USENIX proceedings of the Fourteenth Symposium on Usable Privacy and Security (SOUPS) 2018 [90].

4.1 Research Questions

Our two research questions are:

RQ1: Which of the three studied representations best represents digital aging on social media from a user perspective?

RQ2: What are users' attitudes and concerns relating to digital aging on social media?

4.2 Methodology

Our study explores representations of social media posts to simulate the decay or fading of memories over time. The representations are intended to illustrate that posts are getting older or aging to the viewer. The decay representation is applied to content viewed by "others" as opposed to content that is self-accessed. For example, it is applied to Jane's Facebook profile as viewed by her friends, not content solely viewable by Jane. Aging or decaying of posts has two possible inter-related purposes. It provides temporal context to viewers and it provides some privacy advantages as posts become less accessible by viewers. For this first study, we focus on identifying the best decay representation out of three studied representations from a user perspective, recognizing that further work focusing on the other dimensions will be needed in other studies. Our study also captures users' attitudes and concerns regarding the concept and its potential purposes, including privacy. During the study, we introduced the concept of "decay" as posts getting older over time, but we carefully avoided mentioning "privacy" as a reason why this might be desirable until the very end of the study to avoid unduly influencing participants' perspectives.

To answer RQ1, we gauged users' preferences as determined by responses to Likert-scale questions and interview questions about the preferred representation for use on their own data. Likert-scale questions considered aspects such as meaning, intuitiveness, most natural metaphor, and visual appeal.

To address RQ2, we collected more in-depth answers from users through interview questions and open-ended questions in a wrap-up questionnaire. For example, some questions explored their interpretations and impressions of the representations, if they think the concept of aging/decaying digital artifacts is necessary, and if they would like their own artifacts to age/decay. We also asked about how aging should take place and if they could think of cases in which aging is more useful than deletion or content expiration. Other questions were relevant to the process itself, e.g., what are the thresholds for the aging process, what should the settings look like, and how does this concept relate to their privacy.

The study was cleared by our Research Ethics Board (Clearance ID: 106869). The study methods and questionnaires were pilot tested prior to data collection. We found no major concerns but it helped us slightly refine our study instruments and the flow of the session. Detailed descriptions of the study tasks, interview guide, and questionnaires are available in Appendix B.

4.2.1 Recruitment

Recruitment was done through posters posted across campus and a social media page for advertising the university's HCI studies. Participants were compensated $15 for their time. Before beginning the session, participants read and signed a consent form that explained the purpose and the procedure of the study, and it reminded them that the session was audio-recorded. Personally identifiable information collected from participants was limited to their voice; responses were pseudo-anonymized and non-attributable.

We had 30 participants; 12 were male and 18 were female, with a mean age of 26 (Std. Dev = 9 years). They reported having an average of three social media accounts each and spending an average of three hours (Std. Dev = 2 hours) online daily on social networks. The majority were university students; 16 participants were undergraduate students, 9 were graduate students, and 5 were university staff.

Participants were assigned a username that is not linked to their identity and these usernames were used during data compilation and to report results in the paper. Usernames were generated according to participants' assigned platform (e.g., Facebook: FB1–FB10, Twitter: TW1–TW10, Instagram: IG1–IG10).

4.2.2 Prototype

We created a fictitious social media profile on three different social networks: Facebook, Twitter, and Instagram. We choose several platforms to explore whether our results applied across a range of interfaces. Facebook, Twitter, and Instagram are among the top 5 most popular OSN sites [37, 65] and each has a distinct purpose. We chose platforms that are more likely to have personal content rather than other professional networks such as LinkedIn.

In our prototypes, the profile layout and arrangement imitated the existing look and feel of July 2017 UI versions of each of the three platforms. The content on both Facebook and Twitter was identical; it included miscellaneous photo posts with captions and status updates. To conform with Instagram's layout, its fictitious content included only photos with captions. We intentionally included content that is personal in nature [59], such as family photos, photos of a car with the licence plate number visible, and photos of a house with a visible street address[1]. Status updates included personal sentiments and opinions about potentially sensitive subjects [72] (e.g., political views, support for LGBT).

We implemented decay techniques on the three OSN platforms (Facebook, Twitter, Instragram), using three different approaches: (1) *content fading*, (2) *content pixelation*, (3) *content shrinking*, resulting in $3 \times 3 = 9$ prototypes. Although Novotny's *temporal indices* suggested six other approaches[2] to manipulate older content, we chose the ones that manipulate content consistently and allow for a gradual application. Nevertheless, the other approaches are worth investigating and we address this design choice in Section 4.6.

The dates of the fictitious posts were separated by a month and each prototype showed posts spanning one year. The decay was applied linearly across posts; for example in the fading prototype, transparency levels were reduced by equal increments between any two sequential posts. Figures 4.1 to 4.3 show the nine prototypes on Facebook, Twitter, and Instagram respectively. Each prototype was displayed to the user as a scrollable webpage where they could scroll up and down to view the posts on their own.

Unlike obfuscation techniques from Section 2.5.1, the decay techniques in our prototypes degraded the entire post. To ensure that the limitation posed by those other obfuscation techniques was avoided, we manipulated all the contextual cues related to a post that might be recognizable [78] along with the image itself. These manipulated cues included the image's caption, its intended audience, publishing date and time, comments, date and time of the comments, and tagged friends. Manipulated content was also unclickable to prevent retrieving or accessing the original unmodified post. Moreover, in the shrinking

[1]Prototype photos were downloaded from free websites, 'unsplash' and 'Wikimedia Commons'.

[2]1) Layering newer content on top of older content, 2) moving newer content faster than older content, 3) desatrutating colours and eliminating vowels of older content, 4) changing fashion of content, 5) changing fashion of UI design, and 6) including the profile picture of the user that was used at the time of posting information.

prototypes, zooming in on a browser would not increase the image quality or resolution. This was an intentional design choice to avoid defeating the purpose of the shrinking effect.

4.2.3 Procedure

Thirty participants took part in our 3×3 mixed design lab study featuring one between-subject variable (social media type) and one within-subject variable (decay technique); ten participants were assigned to each of three social media types, and each participant saw all three representations. Assignment of social media types and presentation order of the representations was controlled using a full latin square to ensure that all combinations were cycled and to avoid possible ordering effects. For example Participant X saw {Facebook-Fading, Facebook-Pixelating, Facebook-Shrinking} and Participant Y saw {Twitter-Pixelating, Twitter-Shrinking, Twitter-Fading}.

We collected participants' feedback verbally and through online questionnaires in a 60-minutes session. A session unfolded as follows:

1. View and explore Prototype A
2. Complete representation questionnaire A
3. View and explore Prototype B
4. Complete representation questionnaire B
5. View and explore Prototype C
6. Complete representation questionnaire C
7. Interview/conversation about the concepts and prototypes
8. Complete wrap-up questionnaire

In Steps 1, 3, and 5, participants viewed the social media content as if they were previewing another user's social media profile, not their own. We asked some probing questions while participants viewed each prototype, e.g., what was their interpretation of the representation, what was most appealing/confusing, and whether they would change anything in the design. Other questions explored if the representation was meaningful in terms of conveying the idea that posts were getting old. The prototypes were available to participants if they wished to revisit them during the rest of the session.

In Steps 2, 4, and 6, the representation questionnaires consisted of 10 Likert-scale questions covering whether:

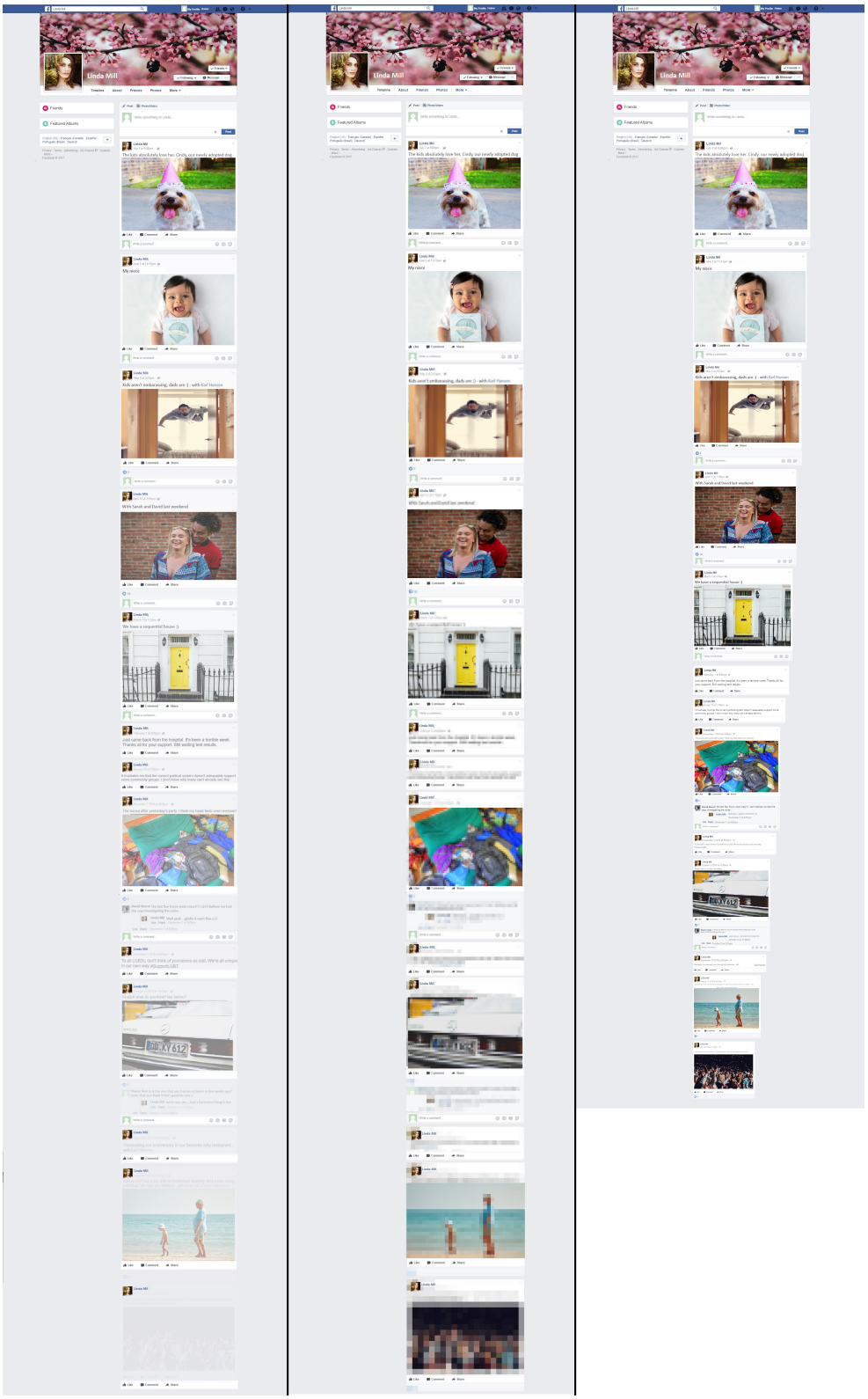


Figure 4.1: Facebook fading (L), pixelating (M), shrinking (R). Posts in the shrinking prototype started the same size as the two other prototypes. The decay effect started at the second post in all 3 conditions.

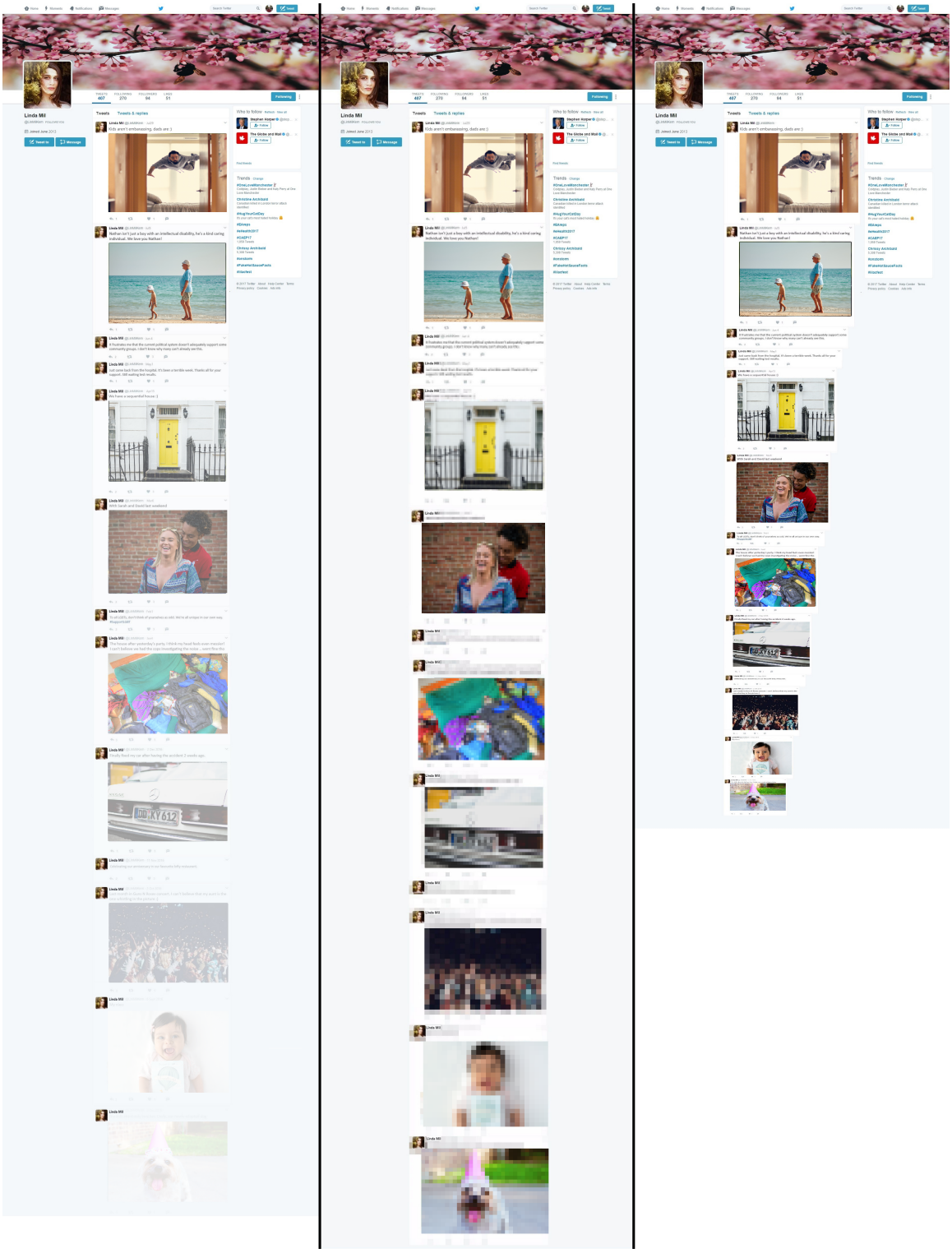


Figure 4.2: Twitter fading (L), pixelating (M), shrinking (R). Posts in the shrinking prototype started the same size as the two other prototypes. The decay effect started at the second post in all 3 conditions.

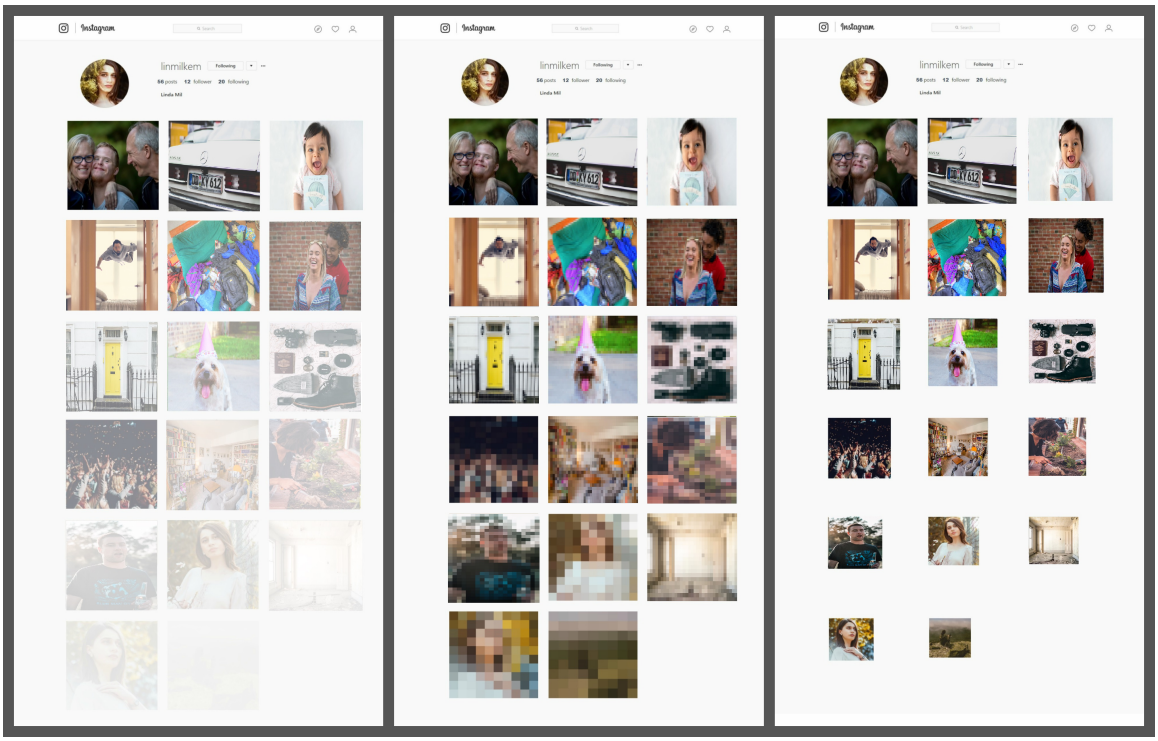


Figure 4.3: Instagram fading (L), pixelating (M), shrinking (R).

- Q1: The visual representation easily showed posts were getting old
- Q2: The visual representation was meaningful
- Q3: The visual representation was confusing
- Q4: The visual representation was complete
- Q5: The visual representation changed their perspective
- Q6: The visual representation was appropriate to the content
- Q7: The visual representation was obtrusive
- Q8: The visual representation of photo posts was intuitive
- Q9: The visual representation of text posts was intuitive
- Q10: They would use the visual representation on their social media account

In Step 7, the wrap-up interview questions sought to learn about users' attitudes and concerns as both *a user browsing another user's profile* and as *an owner of the profile* concerned about other users. For example, we asked for participants' reaction if they came across a profile that uses one of the content decay representations. Other questions examined participants' perception of aging/decaying of digital artifacts, how necessary it is, and by which means it should be implemented in OSNs (e.g., by deletion, expiration, or decay). More questions probed whether participants would use one of the representations to display their own digital artifacts when accessed by other user groups, whether the study changed how they would use social media in the future, and whether content decay would promote their online privacy. In Step 8, the wrap-up questionnaire consisted of one Likert-scale question and three open-ended questions.

In total, each participant gave feedback on three different prototypes, filled out four online questionnaires, and shared opinions pertaining to the concept of content decay.

4.3 Analysis Plan

To answer our research questions, we performed both quantitative and qualitative analyses.

For the statistical analysis, we were primarily concerned with our within-subject variable, *representation type*, with three levels (fading, shrinking, and pixelating). We used Friedman tests (significance level of $p < 0.05$) to test for main effects of representation type. In cases of significant omnibus test results, we followed up with pairwise Wilcoxon signed-rank test with Bonferroni correction applied (significance level of $p < 0.05/3$, i.e., $p < 0.017$).

The qualitative data consisted of audio-recordings from the interviews and open-ended questions from the questionnaires. We transcribed the relevant parts of the interviews. Following content analysis coding methodologies [60], the main researcher compiled the data and extracted the main themes looking for key patterns and particularly insightful feedback through several rounds. A second researcher was involved refining the patterns and interpreting the data, but did not independently code the data.

Table 4.1: Median, mean, and standard deviation (SD) values per question for each approach. All values are out of 5. Highest mean values are highlighted in gray. *Q3 and Q7 were negatively worded; responses were reversed for analysis so that a higher score signifies a more positive response.

Q#	*Fading*			*Pixelating*			*Shrinking*		
	Median	Mean	SD	Median	Mean	SD	Median	Mean	SD
Q1	4.0	3.9	(1.3)	2.5	2.9	(1.5)	4.0	3.5	(1.4)
Q2	4.0	3.6	(1.2)	2.0	2.2	(1.2)	4.0	3.7	(1.1)
Q3*	3.5	3.3	(1.6)	2.0	2.0	(1.2)	4.0	3.3	(1.2)
Q4	3.0	2.8	(1.3)	1.0	1.9	(1.2)	3.0	2.9	(1.3)
Q5	3.0	2.5	(1.2)	2.5	2.5	(1.6)	3.0	2.7	(1.4)
Q6	3.5	3.1	(1.2)	2.0	2.4	(1.4)	4.0	3.3	(1.4)
Q7*	3.0	2.8	(1.3)	1.5	1.9	(1.2)	3.0	3.0	(1.3)
Q8	4.0	3.6	(1.2)	2.0	2.6	(1.5)	4.0	3.7	(1.3)
Q9	3.5	3.3	(1.5)	2.0	2.0	(1.0)	3.0	2.9	(1.5)
Q10	2.0	2.2	(1.4)	1.0	1.6	(1.3)	2.0	2.6	(1.4)

4.4 Results

4.4.1 RQ1 Analysis and Results

We summarize results of our statistical and qualitative analysis pertaining to our first research question: *Which of the three studied representations best represents digital aging on social media?*

Statistical Analysis of Questionnaire

Participants completed ten 5-point Likert scale questions per representation technique. Median, mean, and standard deviation values are available in Table 4.1; higher means indicate more positive scores.

Using the within-subjects variable, representation technique, we compared questionnaire responses to see if participants favoured any technique. We found a significant difference in nine out of the ten questions. Friedman's test results are presented in Table 4.2, with significant differences highlighted in gray. Table 4.3 shows the pairwise comparison between the three approaches and the associated p values (Bonferroni corrected).

Mean responses to the questionnaire ranged from negative to neutral, suggesting that

Table 4.2: Friedman test statistic and significance values. Degrees of freedom = 2, n = 30.

Question	$\chi^2(2)$	***p***
Q1	10.308	0.006
Q2	17.883	0.000
Q3	16.673	0.000
Q4	15.721	0.000
Q5	0.886	0.642
Q6	12.869	0.002
Q7	16.071	0.000
Q8	10.659	0.005
Q9	12.060	0.002
Q10	12.976	0.002

Table 4.3: Asymp. Sig. values as reported from the pairwise comparison using Wilcoxon signed-rank test; Values with Bonferroni-corrected significant differences are highlighted in gray.

Question	***Pixelating-Fading***	***Shrinking-Fading***	***Pixelating-Shrinking***
Q1	0.002	0.350	0.067
Q2	0.000	0.543	0.001
Q3	0.001	0.853	0.000
Q4	0.004	0.792	0.004
Q5	*NA*	*NA*	*NA*
Q6	0.017	0.471	0.017
Q7	0.002	0.388	0.001
Q8	0.002	0.814	0.010
Q9	0.001	0.349	0.015
Q10	0.059	0.169	0.003

participants were generally unenthusiastic about the representation techniques. Reasons for this are discussed in the *feedback on the prototypes* subsection; participants were mainly concerned that the representations might obstruct browsing within social media.

The statistical analysis showed that pixelation was least favourable to participants. Shrinking was the most favourable, but participants did not significantly favour it over fading. Yet, shrinking and fading were significantly more preferable than pixelation.

For completeness, we also verified whether there was a main effect of *social media type* (Facebook, Instagram, Twitter). This was a between-subjects variable and we performed Kruskal-Wallis tests on the 10 questions. We found no significant effect of media type;

with one exception: Q9 showed a significant difference, with Instagram having a lower mean. We believe this single difference occurred because Q9 asked about "text posts", which Instagram does not support.

Feedback on the prototypes

The written and verbal feedback from participants aligned with the Likert scale results: shrinking was the most favourable representation, followed closely by fading; pixelating was least favourable.

As suggested by the feedback for each prototype, detailed next, participants found the shrinking technique most visually pleasing as it looked more "natural". Moreover, it was best associated with memory and the passage of time; putting less significance on older posts by making them tinier. Participants also liked the fading representation because the idea of graying out posts resembled how artifacts fade in real life. In both cases, the representations were reasonable metaphors that provided a logical parallel with their impression of how human memories work. They recognized and brought up their understanding of the metaphors without prompting.

Prototype 1: Pixelating: The initial reaction to pixelation for fourteen participants was that there might be a glitch in the system/website or that the Internet connection was slow and pictures were not loading correctly. Mostly, participants had no idea what was going on. They reported various negative emotions, including thinking of something bad/criminal (FB6), feeling irritated (FB1), angry (IG8), and scared/lost (FB10). In addition, ten participants felt confused or annoyed. Moreover, they thought that someone using the technique on social media must be hiding something from specific people (e.g., non-friends) (n = 7), blocking someone (n = 4), or that the content had been censored (n = 2).

Participants thought that it was pointless to keep posts in such a representation, and felt that it would be better if the post was simply deleted. Overall, participants neither associated such representation with the passage of time nor found it visually appealing. Clearly, the pixelating representation failed to convey the appropriate metaphor, and instead invoked other negative connotations.

Prototype 2: Fading: Fifteen participants found the fading effect intuitive and indicative of its purpose. In addition, it was visually appealing since the gradual fade-out

inherently showed a smoother transition between posts. Eleven participants liked the idea that they could see details about the post, text in particular, compared to the pixelating and the shrinking techniques. Furthermore, the idea of fading the posts resonated for some participants (n = 12) with the metaphor of memories or physical photos fading over time.

As IG3 explained: *"[Fading is] really intuitive, and it's a nice metaphor of fading memories [...] and that's what happens to photos often, when they're older, they get faded [...] but making the pictures smaller? I didn't think of it that way [...] even the pixelated, it was effective, it's visually hard to ignore [...] I just assumed something is wrong with the image [...] so the fading is really nice."*

Nevertheless, some participants (n = 4) thought that faded out content would raise suspicion about the user, for example, suggesting that the user had something to hide. Others were unsure whether they would have guessed its purpose if they suddenly saw this representation on their OSNs.

Prototype 3: Shrinking: Overall, the majority (n = 17) thought the shrinking approach was most intuitive and visually appealing. TW10 explains: *"It's more clever; like fuzzy memories; recent memories occupy more space in your head."*. Participants could see the appeal to instantly realizing what content is most recent without having to look at the dates. As explained by FB3: *"It's like a visual way of seeing that it's a later post [...] the way the time grows the way the grid grows, it kinda correlates that way [...] it would take time to be used to it, but if Facebook had come like this, I'd be more accustomed to it, I wouldn't really have a problem."*

However, some participants (n = 8) initially thought that bigger posts were of higher importance and relevance to the user publishing the content. They believed that the user had somehow chosen to make some posts larger, rather than realizing that size was an automatic characteristic that varied over time. The most common complaint from participants (n = 27) was being unable to have clear legibility of posts as they shrunk. However, between fading and shrinking, they thought shrinking offered better visibility.

Users' preferences: when asked to choose one representation to be applied to their own artifacts, 14 participants favoured the shrinking prototype, 11 participants preferred fading, three participants were undecided between both prototypes, one participant wanted both combined, and one participant preferred pixelation.

RQ1 Summary

Participants expressed clear preferences for the Shrinking and Fading representations, and these successfully conveyed the metaphor of memories fading over time. The Pixelating representation was disliked and held negative connotations; it did not meet the goal of representing aging of digital content. Out of our three evaluated representations, we recommend either Shrinking or Fading as appropriate representations for conveying digital aging/decay.

Interestingly, participants (who were not initially told that this study was about privacy) expressed annoyance resulting from not being able to clearly read the posts as they decayed. Some mentioned a preference for the fading technique because it enabled them to decipher the details of the posts for the longest time. So while they understood the metaphor, they still favoured the representation which showed the least decay. We note that levels of decay could be adjusted for any of these representations and that in an actual implementation, the effect would appear much more gradual since there would likely be more posts in the span of a year. We further discuss the decay implementation formula in Section 7.4.4.

4.4.2 RQ2 Analysis and Results

We next summarize results of our analysis pertaining to answering our second research question: *What are users' attitudes and concerns relating to digital aging on social media?*

We concluded the session with an interview and a wrap-up questionnaire to capture participants' opinions regarding the concept of aging digital artifacts and to discuss if it would increase their online privacy. This part of the session took place after participants had seen all three representations and had provided their feedback about each one. The next subsections summarize the responses from the interview and the open-ended questions of the questionnaire.

Necessity of digital aging

The first question of the wrap-up questionnaire asked *"How necessary is aging of posts in social media?"*. Sixteen participants thought that digital content aging is *very necessary*

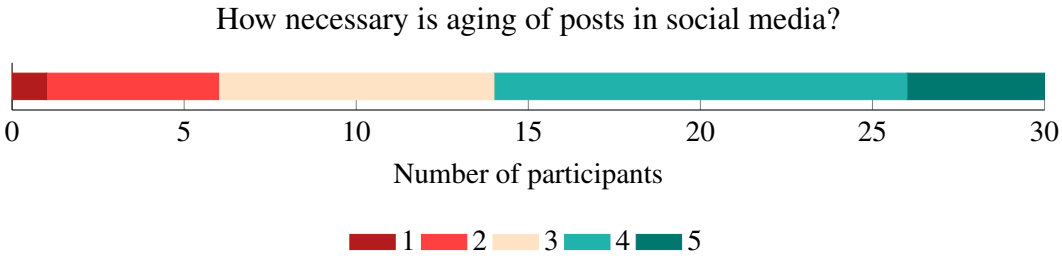


Figure 4.4: Responses to one of the wrap-up questions (1 = Not at all necessary, 5 = Very necessary)

or *necessary*, while eight participants were neutral. Figure 4.4 shows participants' Likert-scale responses.

In our interviews, we asked participants if they would opt-in to the content decay feature for their own content, if available. Two-thirds of participants (n = 20) thought they would, believing that digital content should age, whether to reflect the person they are today, to depict different time periods, or to protect their online privacy. The remaining participants disagreed, or were concerned about how aging of digital artifacts would impact their access to content on their own profiles. While they thought that aging might be appropriate for others viewing their profile, they wanted to retain access to the unedited versions of their own content.

Deletion, expiration, or decay

We further discussed with participants what it means to have their social media content age, and how this should happen.

Eighteen participants recognized that social media content lost relevance as time passed. Two-thirds of participants (n = 20) wanted to either *delete* or *archive* content themselves or potentially have content *decay*. Their choice of method depended on the social media platform and the content itself. Some explicitly mentioned that they wanted to delete content when it no longer reflected their current personalities [51,52,58] or the impression that they wanted to convey to the world.

Secondly, two-thirds of participants (n = 20) saw a need for a *decay* feature on OSNs: one-third would unconditionally opt in and one-third would opt for it conditionally, i.e., if they retained some control over the operation of the decay feature. For example, if it was

programmed to allow an undo of the decay, and if the decay did not apply to their own self-view. Other preferences included being able to select which decay technique should be applied. Moreover, the majority (n = 17) wanted to select which content should decay rather than have it automatically executed. Their choice would depend on specific time thresholds, or the context of the content itself. Fifteen participants thought decay should depend on characteristics of the content more than on how much time has passed. Similarly, eight participants wanted to choose which audience views the decayed content.

Participants (n = 25) thought *decay* would be particularly beneficial in several situations. For instance, they thought it might reduce information overload when browsing other users' profiles. They also thought it would be beneficial if they might regret deletion of specific content. As one participant explained: *"Sometimes you delete something in the spur of the moment then you think I shouldn't have deleted that [...] and there's no point in putting it back cause everybody already saw it [...] with all the comments [...] people regret deleting things." -FB3.* Others thought it would be useful for fact checking data, for keeping track of their online activity, or for archiving or compressing content. As FB4 explained: *"Maybe fact checking data, for politics and election season, sometimes it's important to check news and when they happened, which is something that's easily overlooked in social media."* Others thought it might help to keep only relevant memories and forget irrelevant ones, which might be helpful in the healing process after a breakup. As FB10 illustrated: *"Delete is [...] computation oriented, faded feels more like personal, more human, more like in my memory [...] more natural, I have that association. When you become older, you forget many things [...] right now, social media does not make any differentiation in all our memories, they are all equally relevant, and it happens that along our lives, not all our memories are equally relevant."*

Lastly, two participants thought that the only cases in which content should automatically *expire* is when the person is deceased or the profile is no longer in use. Alternatively, they suggested the family of deceased person could choose to *decay* the content instead.

Privacy

We also wanted to explore participants' perspectives on online privacy. We asked a group of participants[2] if content decay would protect their online privacy. Participants' opinions were polarized. A minority (n = 3) thought the idea does not contribute to online privacy at all. Their main concern was that concealing content would raise questions about the content or the user, hence, they found no contribution to privacy. However, most (n = 11 out of 17) thought it was the only purpose for using decay. For instance, they would decay obsolete content when seeking employment, or when beginning new chapters in their lives. As explained by FB5: *"It would be beneficial to me if I was applying for a new job, or even entering a new relationship, I would not want the company or person to be able to scroll and see my old posts and judge me by them."*

When the eleven participants who thought decay is beneficial for privacy were asked which decay technique is most preferable for privacy, six participants favoured the pixelating technique. They thought pixels hid the content appropriately since pixelation obscures content more quickly by nature. Four participants thought either the fading or shrinking techniques might be helpful to privacy as well, depending on how fast/gradual they decay the content. One participant did not specify a preference.

RQ2 Summary

Participants thought that digital artifacts should age to accommodate changes in their real lives. Decaying digital content was appreciated, and if available on social media, participants would opt-in to the feature. They generally found it useful for online privacy, but responses varied for which representation they would adopt for their own accounts. Specifically for privacy, pixelation was most popular but is also held negative connotations for several participants.

[2]We explicitly asked 17 participants at the very end of the interview. However, it was implicitly discussed with other participants.

4.4.3 Other Comments

Changing of perspective: Eleven participants said that introducing the concept of aging of digital artifacts changed their perspective on how they use social media today. For instance, they intended to go through their own content, re-examine their privacy settings, and re-think which posts remain appropriate for their current lives. This aligns with previous research suggesting that conversations about privacy lead users to reflect on their own practices [2, 64, 111].

We observed a shift during some sessions. Participants initially were concerned about how aging of digital artifacts would affect the visibility of their content to themselves and to others. As the session progressed, they accepted the concept and realized its value for online privacy when displaying content to others.

Other participants (n = 6) expressed no major change in perspective. They were already careful with what they post, or they were accustomed to the look and feel of social media today and saw no reason to change. As one participant noted: *"If I have choice between changing and not changing, I'm not gonna change [...] if they have it changed and I'm forced, I'm not gonna change it either." -FB1*

Downsides of decay: While participants realized that the feature has merit, three participants expressed concerns. Examples in which decay would be problematic include translating decayed content for people with accessibility issues, or when the content is needed as an evidence to verify information (i.e., in a police investigation of a criminal activity[3]).

4.5 Discussion

Our motivation was exploring how to represent the aging of digital artifacts within the UI. We further investigated what aging of digital artifacts means for users and to what extent incorporating this concept within the UI would conform to their sharing and privacy needs. We elaborate on the privacy and design implications of our findings in the following subsections. We then translate those implications into a tentative set of system design recommendations.

[3]We note that this is a misunderstanding on the part of participants; by going through the proper channels, police could still get access to original (now decayed) content.

4.5.1 Aging vs. Privacy Paradox

We found that participants' mental models of how their content should appear online depended largely on whether they were considering aging or privacy at the time. In our study, we intentionally avoided mentioning privacy until late in the session so that we could determine if privacy concerns arose unprompted.

When participants considered the management of their data in terms of aging/decaying, they favoured a gradual fading/shrinking of artifacts over time because it matched their idea that memories lose prominence over time, as suggested by human memory decay theory [17, 118]. As with real memories, they also expected the UI to differentiate between important memories or life events that are clearly remembered despite the passage of time and everyday happenings that are gradually forgotten.

They expressed that the representation should represent the natural forgetting process and should not seem like the artifacts were being manipulated. For example, several participants specifically disliked the pixelation representation because it suggested that something was being intentionally obscured and this raised suspicion.

When prompted to consider privacy implications of digital artifacts, we observed a shift in priorities and requirements. This aligns with previous research regarding the privacy paradox [1,2,5,31,64,95,101,111,126]; people do not intuitively consider privacy risks and sometimes accept them until prompted to consider privacy. Some participants felt that the pixelating representation best reflected the idea of privacy by making it clear that something was intentionally being kept private. Pixelation fit with these participants' mental model of privacy: content was being censored or obscured. They also noted a more discrete dimension to privacy: something should be either kept private or made public. It was not necessarily viewed as a gradual process whereas "aging" was clearly gradual.

We are left with this interesting paradox: users want gradual, natural decaying of digital artifacts (with exceptions for important events) to more accurately reflect human memory, but at the same time want discrete, intentional private/public representation of artifacts to reflect their concept of privacy. For participants, these were two distinct requirements, whereas the literature generally views them as closely related [5, 87, 98, 99].

In both cases, however, participants recognized the benefits of removing irrelevant content and recognized that their preference for the visibility of specific digital artifacts would

likely change over time. The question remains: how do we best reconcile these two distinct intentions while displaying digital artifacts in OSNs?

4.5.2 {Self — Public} Spectrum

Participants require distinct rules when representing aging on their *self* profiles versus their *public* profiles. They wanted their own content to always be visible to themselves, and then had complex rules for how their content should be displayed to different user groups. Those rules differ significantly depending on the category of the content published on their profiles and the intended audience.

Although this was not our intention, participants re-iterated that they expected that the representation of profiles should not be automatically altered to represent aging when *self* accessed. Normally, participants use the web and OSNs as backup repositories to retain their digital possessions [80]. Our participants were concerned that their view of their own data would be altered or the data would become inaccessible without their consent, losing access to the artifacts representing these milestones. Therefore, when the *self* UI visualizes aging/decaying, the default representation should not decay content. While not the intended purpose of decay, the discussion does serve as an anchor for participants' explanation of how things should work for content viewable by others.

When being accessed by the *public/others*, participants desired different rules. Because they are concerned about their online presence and their availability to other online users, it is important that their content is visible to their audience. However, they wish to manage the visibility and aging/decay settings of their online content for both availability and privacy purposes. In this case, the audience comprises a spectrum of *closest friends*, *specific circles of friends*, and moving outwards to the *public*. Participants wished to consider two main factors when visualizing aging on the UI for other audiences: (1) the context/category of the published content and (2) where its intended audience falls on the spectrum. Other practices [78] in the online photo sharing domain similarly adopt a privacy framework by controlling two elements: content and recipient. Indeed the two factors are significant determinants of privacy [78] since some online artifacts are more personal in nature than others [59] (e.g., a self-portrait versus a photo of a landscape). However, the rules are individualized to each user and can be complex as they encompass all possible scenarios and

exceptions. Moreover, rules changed dynamically based on specific contexts or based on exceptions for a specific audience. For example, Joe might enjoy sharing his life memories with others, but Jim prefers having personal photos or embarrassing photos decay when viewed by work colleagues and unmodified when accessed by family members or close friends. Complexity might further increase if Jim also wanted the same artifacts decayed when viewed by a cousin and unaltered for a specific work colleague. Accurately reflecting users' real intentions could quickly become untenable.

This suggests that incorporating controls into the UI that maintain such rules becomes an added effort for users. Firstly, it is impractical that each user can internalize all their desired rules and adjust the rules whithin the UI whenever they publish new content. Secondly, because the desired rules change as time passes and circumstances change, it is unlikely that a system could generalize these rules to match the preferences of every user. This leaves us with another question: should we integrate such complex functionality for controlling the display of digital artifacts in OSNs and can we do so without adding undue effort to users?

4.5.3 Privacy as an intangible subject

The literature show that although users rationally accept privacy risks as a trade-off for the benefits of online sharing, they also express an intuitive concern when prompted [1, 2, 5, 31, 64, 95, 101, 111, 126]. Very few of our participants initially realized the privacy merit of content decay, but opinions evolved throughout the sessions, as presented in Sections 4.4.2 and 4.4.3 (Change in perspective). Initially, participants who favourably viewed content decay said they would opt-in for different purposes. For example, they wanted it as a way of compressing, keeping track of their activity, or forgetting specific memories. Privacy is an intangible subject [64] to users; our participants did not intentionally ignore it, but rather it did not immediately occur to them. However, when prompted about privacy [64] and the ways in which aging of digital artifacts contributes to privacy, they started to realize its potential added value.

In some instances, privacy could be viewed as a positive by-product of decaying content. Some users liked the idea of decaying digital artifacts for reasons other than privacy (e.g., it makes it easier to quickly tell how recently information was posted). These users

might be persuaded to adopt the decay representation due to its perceived usability benefits, but subsequently also gain privacy benefits with no additional effort.

The literature has shown that some Facebook users manage their privacy by trusting their abilities in manually controlling information being shared [1]; few changed Facebook's default privacy settings [50]. Our participants thought they would simply *delete* what they no longer wanted available online. Although they expressed interest in retaining detailed control, practically speaking and, as shown in the literature [2, 50, 126], this is unlikely. Moreover, even if participants had the time and initiative to delete old content, this is actually very difficult to do in OSNs; for example, Facebook only loads a bit of data at a time, in reverse chronological order. And even though the "activity log" allows a user to review older content by year, there is no way to easily access and manage that content. Our participants thought that after the study, they would revise their own OSN content and delete what is no longer relevant. However, this intention only arose because they were specifically primed to consider the privacy of their OSN data [2]. This suggests that normally users remain indifferent to the need to perform retrospective privacy management.

4.6 Limitations

The study had the usual limitation common to lab studies; asking participants to share feedback about a partially unfamiliar concept in a limited amount of time in an artificial environment. A future field study could be designed to complement our findings. Furthermore, the sample size of thirty participants might be small when considering that they were divided across three social media platforms (although every participant saw all three representations), and the university sample of users is not necessarily representative of the whole population.

Additionally, when designing the study prototypes, we distributed fewer than 20 posts across a year to more easily and clearly show the effect of decay. Had we added more posts to the prototypes, the change in representation would have appeared more gradual, which could have impacted participants' opinions. We chose to use artificial data in the prototypes rather than applying the representation to the users' own content. This may have made the content seem more abstract to participants since it was disconnected from any particular context or personal connection. However, this design decision was taken

because protecting the privacy of participants was viewed as more important than the slight methodological advantage to be gained in these early stages of the research.

We evaluated three plausible representations but several other representations exist and could be explored. However, we focus on three to avoid participant fatigue. We chose to test the decay technique as a within-subject variable to enable each participant to compare the different techniques against each other. Platform was tested between-subjects to reduce the number of prototypes each participant needed to visit. We anticipated that this combination would give participants the best overview within the constraints of our session length. Other study designs may yield different insights.

This research has led to several possible future research directions. The design of a future study could consider scenarios to help users with specific contexts. Another future study could empirically examine how aging/decaying digital artifacts on an OSN profile affects viewers' impression of its owner. This could be explored in several different social contexts: political, employment, or relationships/dating contexts.

4.7 Conclusion

This chapter presented a lab user study exploring the concept of aging or decaying of digital artifacts and reported results from both qualitative and quantitative analyses. Results showed an inclination towards representations that closely represent fading memories over time. Because of the nature of human memory, and users' mental model of privacy, we identified distinct user requirements when addressing either aging or privacy in the UI. These two distinct purposes should be further explored to determine how they can be best reconciled in interaction design.

A balanced approach to addressing users' requirements would seek to promote privacy while minimizing user effort and simultaneously enabling user reflection. Towards this goal, we provided three preliminary design recommendations. Although decay features do not address every aspect of online privacy and long-term data availability dimensions, it can help minimize the potential unintended consequences associated with data availability on OSNs.

To summarize, this work compares three OSN content decay representations, investigates users' attitudes and concerns about the aging of digital artifacts, and provides early recommendations that would contribute to users' privacy and sharing needs. We also believe the study is a step towards answering currently open research questions pertaining to visualizing passage of time in OSNs.

Chapter 5

The Influence of Decay Representations on Hiring Decisions

Bringing together Chapters 3 and 4, we evaluate how *decay representations* influence online reputations in a hiring context. Results from the previous chapter look promising as users found both the *fading* and *shrinking* representations to be intuitive and successful at representing the metaphor of decayed memories [90]. However, how such representations enhance online privacy in situations that involve a critical assessment by the viewer (e.g., managers looking at a job candidate's profile) needs to be investigated. We chose the hiring context rather than political because it has potential for directly impacting more users.

Through an online crowdsourced survey, we showed managers fictitious OSN profiles for job applicants. We explore whether the use of *decay representations that apply shrinking* influence managers' opinion of job candidates and impact their hiring decisions. We also examine whether the gender of the profile owner and managers' demographics such as gender and age impact these decisions. We further probe how online reputation shapes their real life hiring decisions.

5.1 Research Questions

Our research questions are:

RQ1: How does decaying a candidate's OSN profile influence managers' hiring decisions?

RQ2: How does the gender of the candidate influence managers' hiring decisions?

RQ3: How do managers' demographics such as gender and age influence their hiring decisions?

RQ4: How do OSNs and online reputation influence managers' real-life hiring decisions?

5.2 Methodology

To address our research questions, we conducted a 3×2 between-subjects study with real managers, with the applicant's gender (male, female) and type of representation (original, decayed, control) as the independent variables. The *decayed* representation showed a Facebook profile with posts gradually shrinking in size. The *original* representation showed a Facebook profile with all posts full size. The *control* representation showed a Facebook profile with no posts displayed on the timeline.

We explore the effect of the applicant's gender, by having parallel *male* and *female* profiles. Other attributes such as age, ethnicity/race, or cultural background were left out of our current study design. We further address this design choice in both the discussion and limitation sections (Sections 5.5.2 and 5.6 respectively).

Given that we found no effect of social media platform in Chapter 4, we chose to test only Facebook in this study since previous literature [34] says that it is often used during hiring processes. We also chose Facebook as a platform with personal content rather than a platform with professional content such as LinkedIn. LinkedIn profiles mainly display professional information that extends a résumé, but we wanted to explore the decay effect on content that is personal in nature. We prototyped six fictitious Facebook profiles: female-decayed (FD), female-original (FO), male-decayed (MD), male-original (MO), female-control (FC), male-control (MC). In our scenario, the owner of each profile was applying for a job.

We designed an online crowdsourcing survey where participants received a link to a video displaying one profile corresponding to their study condition. Participants were randomly assigned to one of six conditions. We launched small batches of the survey in parallel (to even out the number of male/female managers viewing the survey link).

We answer RQ1 through Likert-scale questions that collected managers' hiring decision and asked them to evaluate aspects of the candidate's character and online reputation. We also look at what influenced their decisions through an open-ended question.

We answer RQ2 by analyzing if the gender of the *candidate* impacted responses.

We answer RQ3 by testing if *managers'* gender and age had an impact on the responses.

We answer RQ4 through Likert-scale questions that explored managers' likelihood to search job candidates online or hire them despite of having previously published negative

content. The questions further considered the timeframe of past online activities managers consider relevant, and identified the most important factor in their hiring decision (online reputation, skills and qualifications, or both equally). We further look at how social media shapes their decision through an open-ended question.

The study was cleared by our Research Ethics Board (Clearance ID: 109257). The study was pilot tested with four participants. Three of four pilots were run with individuals who had performed managerial or recruitment duties on behalf of their employers.

5.2.1 Prototypes and Videos

Using Mockplus[1] 3.3.2.4, we prototyped six Facebook profiles containing fictitious data. The layout matched the look and feel of the Facebook UI as of June 2018. The control profiles (FC, MC) showed no posts on the timeline, to appear as if the user had restricted access to non-friends and made all posts private. The content of the four prototypes with visible posts was identical except for small changes to match the gender of the owner, such as the user's name, profile picture, and image uploads[2]. When we needed to vary the images, we used images that closely mirrored each other, except for the gender of the person in the photo. Each profile had 26 posts in total. Posts' dates were separated by 2-10 months, and spanned 11 years (2018 to 2007). We included content that indicates positive, neutral, and negative aspects (discussed in Section 2.3.1) of the profile owner's personality, activities, and his/her online reputation. We included posts that reflect general day-to-day activities or might show a positive attitude or lifestyle, i.e., posts of the user doing outdoor activities, expressing gratitude, and socializing with friends and family. We also included posts that might suggest negative connotations. We chose three types of online content that have previously led to rejection among employers [34, 73, 89, 133]: drinking alcohol and comments showing irresponsible behaviour, offending a past employer, and partaking in a controversial activity. Positive, neutral, and negative content was interspersed throughout the profile.

The prototypes were presented to our participants as videos. We created a separate HD video for each prototype[3]. The videos showed the photos/friends page (with text stating

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

[2] Prototype photos were downloaded from free websites, 'unsplash' and 'Wikimedia Commons'.

[3] The six videos can be accessed at:

that the friendlist is only visible to the user's friends) and slowly scrolled through the user's content published on the timeline. The length of the videos was between 3:14 and 3:34 minutes. The videos let the managers view the content (and pause/move back/forth) without directly interacting with the fake profile (which was not fully functional).

The original condition: Posts of the Original condition had equal size across the timeline. All posts were displayed and none were compressed by year[4]. The size of images and the textual attributes of the post (e.g, captions, date, usernames, comments, tags) were not manipulated.

The decayed condition: In the previous chapter, we recommended *shrinking* content as a representation to represent aging of digital content. In such representation, Facebook posts on a user's timeline become gradually smaller and less legible as they become older. We implemented the shrinking representations in the two prototypes that *decay* content. The *shrinking* rate was applied linearly across posts from completely undecayed to illegible because of its small size. All posts gradually shrunk over time at the same rate; shrinking encompassed images and all textual attributes (e.g, captions, date, usernames, comments, tags). In a full implementation, clicking on the posts would not have returned them to their full size (except if accessed by the profile owner). Similarly, shrunk posts would not have improved resolution if the viewer tried to zoom in on a browser.

Content that showed alcohol posts and comments indicating irresponsible behaviour were much deteriorated in the decayed condition because they were old posts on the timeline. Other older posts with positive or neutral behaviour were also deteriorated (e.g., a picture of a device developed in a course project, or reminiscing over university). Posts showing controversial activities (e.g., participating in a protest and sharing political opinions) or an offence to a past employer were more legible as they were newer posts.

Figure 5.1 shows a partial snapshot of the older posts on the male and female profiles,

---

`https://youtu.be/j3BHhfC7tMU` *(FD)*,

`https://youtu.be/mB31oZHy5kM` *(FO)*,

`https://youtu.be/WVl5cD7AJ_0` *(MD)*,

`https://youtu.be/49p7r273awI` *(MO)*,

`https://youtu.be/6ojnVz89G3M` *(FC)*,

`https://youtu.be/q4lB9fYbsis` *(MC)*

[4]Facebook used to compress posts by year, where a clickable bar with a year label was displayed instead of posts. The posts were only displayed when the user clicked the bar to expand them.

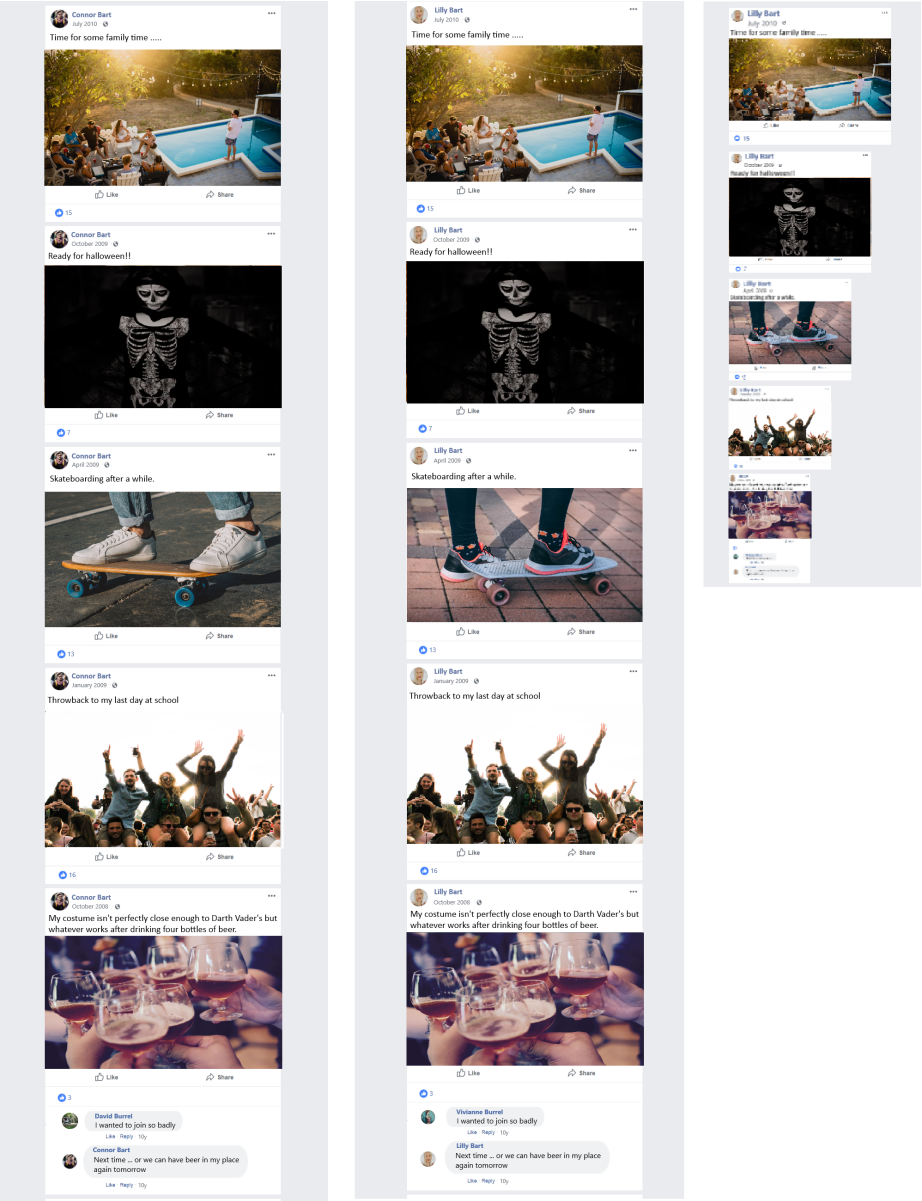


Figure 5.1: The same partial content in 3 study conditions. L: *male-original* condition (MO). M: *female-original* condition (FO). R: *female-decayed* condition (FD). Posts started the same size at the top of the profile for both the Original and Decayed conditions. This image illustrates posts further down the timeline, showing their smaller size in the decayed condition.

and how it was displayed on the female decayed prototype. The selected content reflects older posts (towards the end of the profile and video), illustrating how content has shrunk over time in the decayed condition.

The control condition: The two control profiles with no posts were identical expect for the user's name and the profile picture. On Facebook, if a user decides to limit access to all of their past history, the user's timeline becomes blank (only a dot is displayed instead of posts) when accessed by non-friends. The profile picture and personal information on the left hand-side of the profile can be displayed as configured by the user. Our control profiles showed no posts and no personal information, as shown in Figure 5.2.

We note that there could be other variations for the control profile. For example, it could have shown only a few posts. However, this approach has several issues. It is unclear what the cut-off should be, particularly because it would be impractical to decide at which point the post becomes illegible in the equivalent decayed profile. In practice, Facebook currently provides the option for users to turn all their past posts to private. Users in real life who are revising their privacy settings are more likely to choose this option rather than retroactively adjust the privacy setting [101,152] of individual posts. Consequently, a blank control profile seemed to be the most straightforward choice.

Candidates' Persona: Our fictitious candidates are Caucasian, in their late-twenties, and have graduated from electrical engineering. They have a work experience at Ericsson and are currently seeking employment. Their interests include traveling, skateboarding, and spending time with family and friends.

5.2.2 Questionnaire

We had a total of 75 questions covering 7 categories (A-G) described below. The full questionnaire is also available in Appendix C. Managers completed Sections A and B, watched the video, then completed the remaining sections. After the video, we told the managers to assume that they had already interviewed the candidate and determined that they had the required skills and qualifications. They subsequently visited the applicant's Facebook profile and were deciding whether to hire them.

A. Demographics: We collected participants' gender, age, highest level of education, field of expertise, years of work experience, example duties for their job, average number of hiring decisions they make per year, and familiarity with OSNs.

B. Ratings: We gauged participants' perception of desirable traits for job candidates in general. Participants rated ten traits using 5-point Likert-scale questions ('extremely

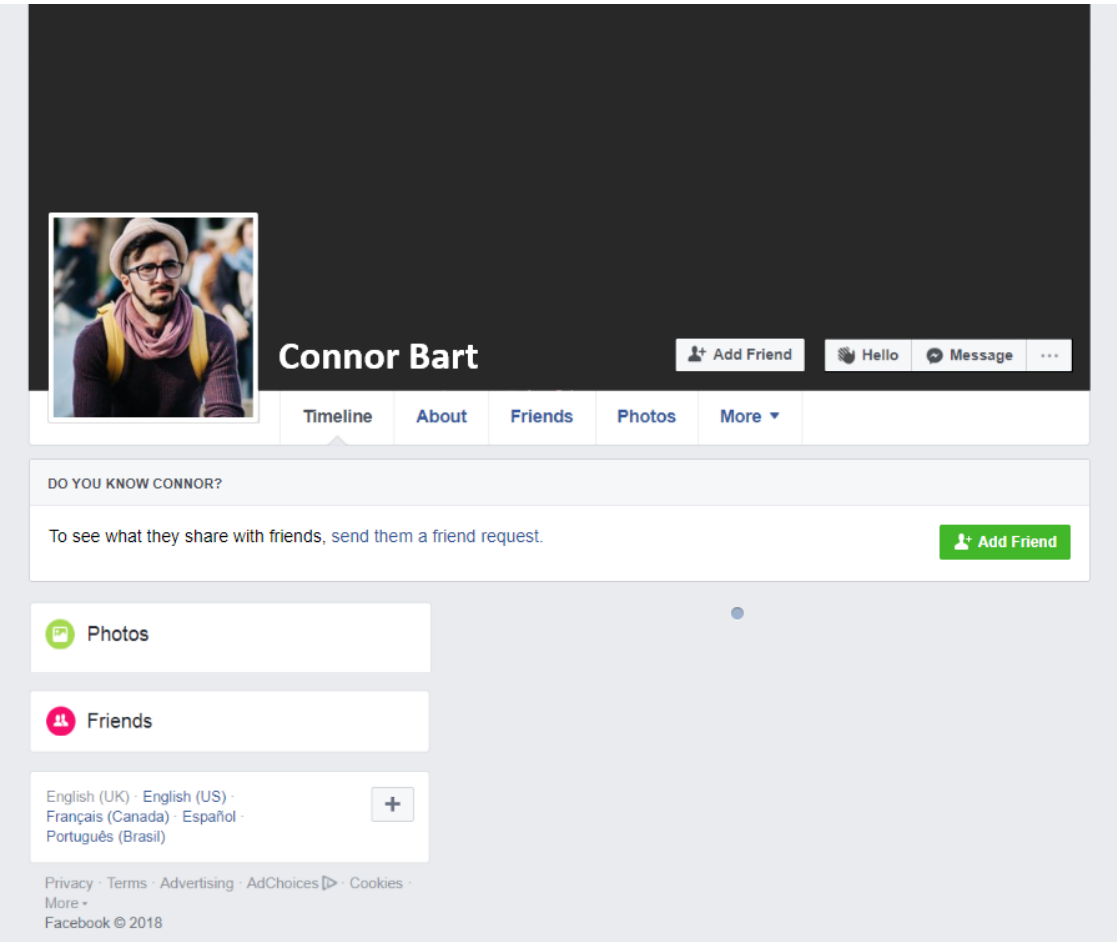


Figure 5.2: The profile of the male candidate in the *Male Control (MC) condition*.

important' to 'not at all important'). These traits had been identified as desirable in previous literature [19, 97, 106]. The traits were: *(1) hard working*, *(2) self-motivated*, *(3) loyal*, *(4) dependable*, *(5) team-oriented*, *(6) confident*, *(7) adept communicator*, *(8) respectful to work ethics*, *(9) flexible*, *(10) cautious*.

As discussed in Section 2.3.1, the literature has identified aspects of online reputation that influence managers' hiring decision. These aspects included lifestyle [34, 73, 89], personal portrayal [34], professionalism [133], appropriateness of content [73, 89, 133], use of alcohol or drugs [73], and comments about an employer [73, 91]. We used these characteristics in a set of 10 rating questions relating to online reputation. While these cover a range of characteristics, this list is not exhaustive and other aspects may exist. Using

5-point Likert-scale questions, we asked participants to rate how important it is that a candidate's online reputation *shows* each of the following five *desirable*[5] aspects: *(1) positive lifestyle*, *(2) positive attitude*, *(3) professionalism*, *(4) volunteering and charitable giving*, *(5) openness and willingness to experience*.

We followed with how important it is that a candidate's online reputation *avoids* each of the five *undesirable*[5] aspects: *(1) a history of using drugs*, *(2) a history of drinking alcohol*, *(3) offence or criticism to past employer or coworker*, *(4) inappropriate content*, *(5) controversial content*.

C. Video verification: We included five questions about specific content seen in the video to ensure that our managers have played and watched the video. These questions asked about content from the start, middle, and end of the video. After answering this section, the video was made available to the managers for reference when answering Sections D, E, and F of the questionnaire.

D. Hiring decision: We asked questions relating to hiring the candidate after managers had viewed the candidate's profile in the video. Using 5-point Likert-scale questions, we asked how likely they were to hire the candidate, and how likely they would be to do more research about the candidate before hiring them. We also asked an open-ended question to understand what influenced their hiring decision.

E. Candidate evaluation: To explore our managers' opinions of the candidate they saw in the video, we asked them to evaluate the candidate against the personality traits (ten 5-point Likert-scale questions) and aspects of online reputation (ten 5-point Likert-scale questions) from Section B. In total, we had twenty 5-point Likert-scale questions for managers to assess the candidate.

F. Visual representation and privacy: We asked if it was easy for managers to see that posts were getting older and to differentiate between older and more recent posts. We also asked if the job candidate values his/her online privacy. These questions used a 5-point Likert-scale. Furthermore, we asked whether the layout of posts and how they were displayed had a positive influence, a negative influence, neither, or no influence on the hiring decision.

G. Real-life practices: We asked about managers' own real-life practices as people

[5]We did not state in the survey that these were *desirable* or *undesirable*.

responsible for making hiring decisions. We asked if they look up candidates online before hiring them and to list social media sites they check before hiring decisions We included an open-ended question to capture in more detail how online reputation influences their hiring decision. We also asked them to identify the most important factor in their hiring decision: online reputation, skills and qualifications, or both equally. Similar to work presented in Chapter 3, we examine how likely our managers were to hire a real-life job candidate whose online activities: (1) could damage the company's reputation, (2) are against the company's values, (3) show participation in illegal activities, (4) show inappropriate content, (5) show controversial content, (6) show negative comments about past employers. These were 5-point Likert-scale questions. We also explored what timeframe of online content our managers would consider relevant when making a hiring decision.

5.2.3 Recruitment and Participants

We used Qualtrics[6] and TurkPrime[7] to administer our surveys. TurkPrime is an internet-based platform that facilitates online crowdsourcing recruitment through Mechanical Turk. TurkPrime workers are ranked according to their previous history at accurately completing tasks and giving responses accepted by requesters (namely, approval rate). We included workers who have an approval rate of at least 97%. We also required workers with specific demographics: (1) aged between 25 and 65 years old[8], (2) employed in managerial positions, (3) residing within the USA. TurkPrime provides advanced filtering criteria and they only launched the survey to workers who pre-identified as "managers" from a long list of occupations. Responses to the open-ended question about work duties performed indicate that our managers were in administrative roles including managing, recruiting, or leading employees. We note that "manager" is a broad category, and may not necessarily reflect users in high-paying executive positions.

[6] `https://www.qualtrics.com/`

[7] `https://www.turkprime.com/`

[8] Although we configured TurkPrime to recruit managers aged between 25 and 65 years, we got three responses who reported being 23 and 24 in our demographics questionnaire.

The actual time to complete the surveys was 17 minutes (on average). Our participants were paid a fixed fee following the current US federal minimum wage and received USD$2.25. With administration fees, we paid a total of USD$3.59 per participant. TurkPrime had the option to exclude workers who have completed previous surveys. Therefore, we collected unique responses only.

Data collection took place between August 2018 and January 2019. We collected a total of 369 responses. After validation of the verification questions and handling of irregular responses, we had a total of 360 valid responses.

Most managers had at least a 4-year college degree. They were also "very familiar" with social media. They came from 40 various fields, but most were in business, finance, manufacturing & production, the service sector, and technology & computers. Table 5.1 summarizes managers' demographics. Only eight out of 360 participants reported an average of 0 hiring decisions per year. The median values per condition were relatively similar: 6.5 decisions in FD, 4.0 in FC, and 5.0 in the other four conditions. The very large maximum values from Table 5.1 were reported by a small handful of participants. We note that the total number of managers is lower in some conditions because we reached a saturation point where no more managers were participating, combined with having eliminated invalid responses.

5.3 Analysis Plan

Statistical analysis was performed using SPSS 25, with significance value set at $p < 0.05$. For analysis of Likert-scale questions, a score of 1 was assigned to the most negative response (i.e., managers were less accepting of the candidate) and 5 to the most positive response (i.e., managers were more accepting of the candidate). For our first three research questions, we focus on reporting results to the main questions relating to the hiring decisions: the hiring decision question (*V1-HIR*), the open-ended question from Section D, and all 20 questions from Section E. These questions primarily prompted our managers to evaluate their respective job candidate. We also report results of the open-ended question from Section G, relating to real-life practices.

The 20 questions from Section E represent three compound measures: evaluation of the candidate's *personality traits*, *positive aspects of online reputation*, and *negative aspects*

Table 5.1: Managers' demographics per condition.

	Category	*FD*	*FO*	*MD*	*MO*	*FC*	*MC*
Gender	Male	38	33	31	28	26	24
	Female	30	27	30	34	29	29
	Other	-	1	-	-	-	-
	Total	68	61	61	62	55	53
Age	Minimum	23	25	23	28	26	25
	Maximum	63	64	62	63	64	63
	Mean	39	39	39	39	40	41
	Std Dev.	10	10	10	8	9	10
Number of hiring decisions	Minimum	0	0	0	0	0	0
	Maximum	200	500	300	140	100	30
	Mean	19	26	15	18	9	8
	Std Dev.	36	78	39	29	15	7
Years of work experience	Minimum	1	1	1	1	1	0[1]
	Maximum	23	33	30	25	40	31
	Mean	9	10	9	9	9	8
	Std Dev.	6	7	7	6	7	6

[1] Less than one year.

of online reputation. Each measure included at least five Likert-type ordinal questions. Following Boone and Boone [15]'s methodology, the individual questions in each measure were combined by calculating a mean/composite score before running statistical tests. We calculated a mean score of the ten questions evaluating the candidate's personality traits (*V2-PER*), a mean score of the five questions asking about positive online reputation aspects (*V3-POS*), and a mean score of the five questions that considered negative online reputation aspects (*V4-NEG*). These compound variables were used in statistical analyses instead of the individual questions. Table 5.2 shows descriptive statistics of the hiring question the 20 questions constituting the three compound variables.

To answer RQ1–RQ3, we ran statistical analysis on four variables: V1-HIR (hiring decision, ordinal Likert-scale data), and the three compound variables V2-PER, V3-POS, and V4-NEG (continuous data). Depending on the research question, and detailed in the following sections, we ran one-way ANOVA on V1-HIR to V4-NEG responses to look for main effects, and we used two-way ANOVA on V1-HIR to V4-NEG to look for interaction effects. We recognize that opinions diverge around the use of parametric tests (one-way and two-way ANOVA) with non-parametric data (V1-HIR). Previous work similarly used parametric tests with ordinal data [66, 96]. As suggested by Norman [96], parametric tests

Table 5.2: Likert-scale responses for the hiring decision and the 20 individual questions used in the compound variables (V2-PER, V3-POS, V4-NEG), organized by either *representation* or *gender*, with higher means in bold red.

Q#	*Representation*									*Gender*					
	Decayed (n= 129)			*Original (n= 123)*			*Control (n= 108)*			*Female (n= 179)*			*Male (n= 180)*		
	Md	μ	SD	Md	μ	SD	Md	μ	SD	Md	μ	SD	Md	μ	SD
D1: Hiring	4.0	3.8	(1.0)	4.0	3.3	(1.3)	4.0	**4.0**	(1.0)	4.0	3.7	(1.1)	4.0	3.7	(1.2)
E1: Hard working	4.0	**3.6**	(1.0)	3.0	3.1	(1.1)	3.0	2.6	(0.8)	3.0	3.1	(1.0)	3.0	**3.2**	(1.1)
E2: Self-motivated	4.0	**4.1**	(0.9)	4.0	3.6	(1.1)	3.0	2.7	(0.9)	3.0	3.4	(1.1)	4.0	**3.6**	(1.2)
E3: Loyal	4.0	**3.7**	(1.1)	3.0	3.2	(1.2)	3.0	2.7	(0.9)	3.0	3.2	(1.1)	3.0	3.2	(1.1)
E4: Dependable	4.0	**3.6**	(1.1)	3.0	2.9	(1.2)	3.0	2.7	(0.9)	3.0	3.1	(1.1)	3.0	3.1	(1.1)
E5: Team-oriented	4.0	**3.8**	(0.9)	4.0	3.5	(1.0)	3.0	2.6	(0.8)	3.0	3.3	(1.1)	3.0	**3.4**	(1.0)
E6: Confident	4.0	**4.3**	(0.8)	4.0	4.1	(1.0)	3.0	1.0	(0.9)	4.0	3.8	(1.1)	4.0	3.8	(1.1)
E7: Adept communicator	4.0	**4.0**	(0.9)	4.0	3.6	(1.1)	3.0	2.7	(0.9)	4.0	**3.5**	(1.1)	3.0	3.4	(1.1)
E8: Respectful to work ethics	3.0	**3.4**	(1.1)	3.0	2.7	(1.2)	3.0	2.7	(0.9)	3.0	3.0	(1.2)	3.0	3.0	(1.2)
E9: Flexible	4.0	**3.9**	(0.9)	3.0	3.2	(1.0)	3.0	2.8	(1.0)	3.0	3.3	(1.0)	3.0	3.3	(1.0)
E10: Cautious	3.0	**3.2**	(1.1)	2.0	2.3	(1.0)	3.0	2.7	(0.8)	3.0	3.0	(1.2)	3.0	3.0	(1.2)
E11: Positive lifestyle	5.0	**4.4**	(0.7)	4.0	3.7	(1.2)	3.0	3.5	(1.2)	4.0	**3.8**	(1.1)	4.0	3.7	(1.2)
E12: Positive attitude	5.0	**4.3**	(0.9)	4.0	3.7	(1.2)	3.0	2.8	(0.9)	4.0	**3.7**	(1.1)	4.0	3.6	(1.2)
E13: Professionalism	4.0	**3.6**	(1.2)	3.0	2.9	(1.2)	3.0	2.9	(0.9)	3.0	3.2	(1.2)	3.0	**3.3**	(1.2)
E14: Volunteering	3.0	**3.3**	(1.1)	3.0	2.7	(1.1)	3.0	3.2	(1.1)	3.0	**2.9**	(1.1)	3.0	2.8	(1.1)
E15: Openness	4.0	**4.2**	(0.8)	4.0	4.0	(1.0)	3.0	2.6	(0.9)	4.0	3.7	(1.2)	4.0	3.7	(1.2)
E16: History of drugs	5.0	**4.3**	(1.0)	4.0	4.1	(1.0)	4.0	4.0	(1.0)	5.0	4.1	(1.0)	5.0	**4.2**	(1.0)
E17: History of alcohol	4.0	3.8	(1.2)	2.0	2.2	(1.4)	4.0	**4.0**	(1.0)	3.0	3.3	(1.4)	3.0	**3.4**	(1.4)
E18: Offence to employer	3.0	3.2	(1.6)	2.0	2.0	(1.3)	4.0	**4.0**	(1.0)	3.0	3.0	(1.6)	3.0	3.0	(1.5)
E19: Inappropriate content	4.0	4.0	(1.1)	3.0	3.2	(1.2)	5.0	**4.1**	(1.0)	4.0	3.7	(1.2)	4.0	**3.9**	(1.1)
E20: Controversial content	3.0	3.4	(1.3)	2.0	2.7	(1.3)	5.0	**4.1**	(1.0)	3.0	3.4	(1.3)	3.0	3.4	(1.3)

are robust in regards with violations posed by non-parametric data. We believe that one-way and two-way ANOVA tests seemed to be our best option.

For the two open-ended questions, we used *an inductive* content analysis approach [38] to look for main themes across responses. Two researchers extracted the main themes and looked for key patterns through several rounds. A third researcher was involved in handling unclear cases and to better refine the coding.

5.4 Results

We first present our managers' general preferences when it came to the 20 traits for Section B of the questionnaire (answered before watching the video). Figure 5.3 shows the 10 personality traits and their importance to our managers. Figure 5.4 similarly shows the importance of the 10 aspects of a candidate's online reputation. Overall, managers report being more interested in a candidate's personality than his/her online reputation and activities.

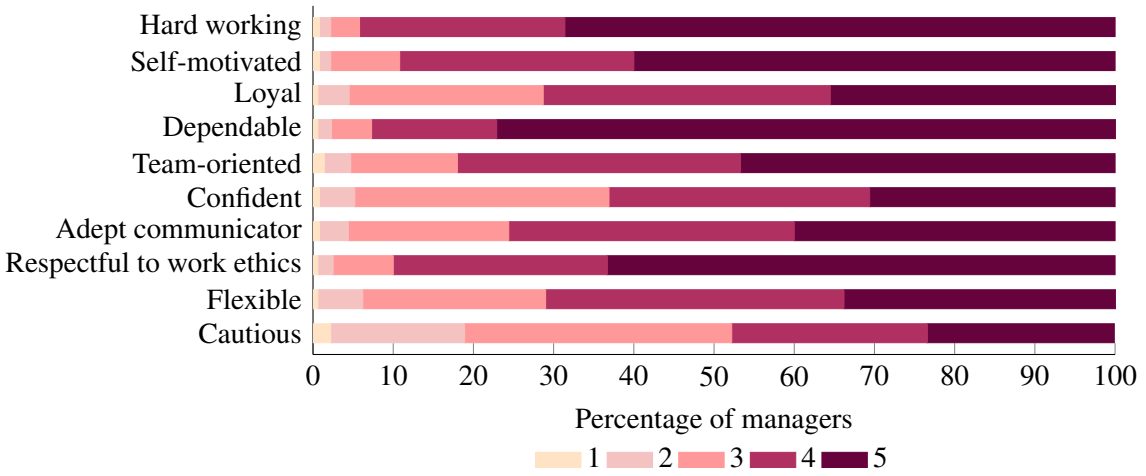


Figure 5.3: Ten personality traits as rated by managers (1 = Not at all important, 5 = Extremely important).

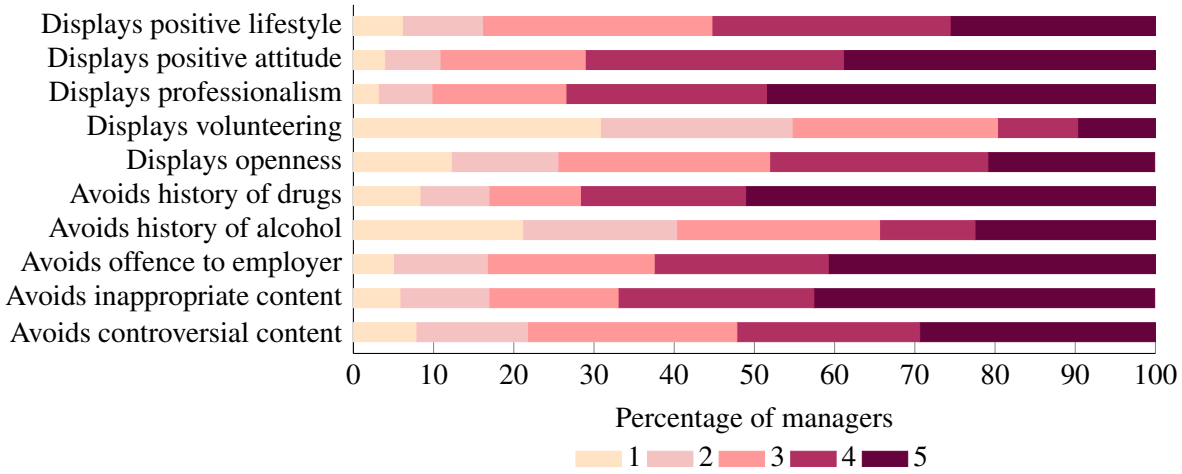


Figure 5.4: Ten aspects of online reputation as rated by managers (1 = Not at all important, 5 = Extremely important).

5.4.1 RQ1 Analysis and Results

RQ1: *How does decaying a candidate's OSN profile influence managers' hiring decisions?*

To answer our first research question, we ran one-way ANOVA on *V1-HIR*, *V2-PER*, *V3-POS*, and *V4-NEG* to look for main effects of our independent variable, *representation type*.

The hiring decision made by managers in each of our three study conditions is shown in Figure 5.5. Managers of both the decayed (71%) and the control (70%) representations

Table 5.3: RQ1: Effect of representation type on each of the 4 dependent variables. The table shows mean and standard deviation (SD), one-way ANOVA test results, and Games-Howell pairwise results. The highest means are in bold red and significant results are highlighted in gray. F: f-distribution, p: significance value, Partial η^2: effect size, D: Decayed, C: Control, O: Original representation.

	Mean (SD)			*ANOVA*			*Pairwise Comparison*		
V#	D	C	O	F(2, 357)	p	Partial η^2	D - O	D - C	O - C
V1-HIR	3.8 (1.0)	**4.0 (1.0)**	3.3 (1.3)	11.254	<0.0005	0.059	<0.006	0.253	<0.0005
V2-PER	**3.7 (0.7)**	2.8 (0.8)	3.2 (0.8)	44.704	<0.0005	0.200	<0.0005	<0.0005	<0.0005
V3-POS	**4.0 (0.6)**	2.8 (0.8)	3.4 (0.9)	67.535	<0.0005	0.274	<0.0005	<0.0005	<0.0005
V4-NEG	3.7 (1.0)	**4.0 (0.9)**	2.8 (0.9)	56.081	<0.0005	0.239	<0.0005	0.046	<0.0005

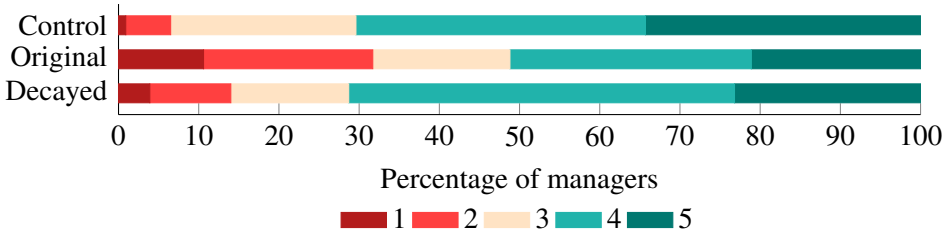


Figure 5.5: Likert-scale responses to the hiring decision question as provided by our managers per condition. 1= Extremely unlikely, 5= Extremely likely.

were either somewhat or extremely likely to hire the candidate compared to those (51%) who saw the original representation.

V1-HIR: The one-way ANOVA test[9] showed a significant effect of representation on V1-HIR, $p < 0.00005$. We followed up with Games-Howell post-hoc analysis for pairwise comparisons. We found a significant difference between the Original–Control pair ($p = 0.000$) and the Original–Decayed pair ($p = 0.006$), but not between Control–Decayed. The control ($\mu = 4.0$, standard deviation $(SD) = 1.0$) and decayed ($\mu = 3.8, SD = 1.0$) representations most positively influenced managers' decision, followed by the original representation ($\mu = 3.3, SD = 1.3$).

V2-PER - V4-NEG: We found a significant effect of representation on all three variables ($p < 0.0005$). We followed up with Games-Howell post-hoc analysis for pairwise

[9]We also ran independent samples Kruskal-Wallis (KW) test to look for main effects. KW showed the same reported main effects of the one-way ANOVA test. We found a significant effect of representation on V1-HIR, $p = 0.000$. Pairwise comparisons (Bonferroni adjusted) showed a significant difference between the Original–Control pair ($p = 0.000$) and the Original–Decayed pair ($p = 0.017$), but not between Control–Decayed.

Table 5.4: RQ1: Key factors influencing managers' hiring decision in the Original and Decayed conditions. Responses indicating a positive assessment are in green and negative assessments are in red. Colour intensity corresponds to the response's popularity.

	Positive Personality Traits	Positive Lifestyle	Has Matured	Qualifications & Intellect	No Red Flags	Unprofessionalism	Politically Active	Partying Lifestyle
FO (n= 61)	8%	7%	16%	30%	7%	61%	10%	10%
MO (n= 62)	29%	6%	19%	23%	10%	50%	16%	21%
FD (n= 68)	60%	24%	6%	26%	19%	29%	9%	-
MD (n= 61)	39%	20%	5%	23%	21%	20%	7%	-

Table 5.5: RQ1: Key factors influencing managers' hiring decision in the Control condition. Responses indicating a positive assessment are in green and negative assessments are in red. Colour intensity corresponds to the response's popularity.

	Qualifications & Interview	Privacy	Nothing Negative	No Enough Information to Make a Judgment	Assumptions About User	Questioning Lack of Content
FC (n= 55)	42%	31%	18%	25%	15%	7%
MC (n= 53)	47%	21%	26%	23%	17%	11%

comparisons. Managers who saw the Decayed representation (D) gave the most positive responses when evaluating the candidate's *personal traits* (V2-PER) and his/her *positive aspects of online reputation* (V3-POS) compared to the Original (O) and Control (C) representations, with statistically significant results for all pairwise comparisons ($D > O > C, p < 0.0005$). On the other hand, managers who saw the Control representation gave the most positive responses when evaluating the candidate's *negative online reputation* (V4-NEG), a statistically significant result between the Decayed–Original and the Control–Original pairs ($(C \approx D) > O, p < 0.0005$). Table 5.3 shows means and standard deviation, ANOVA test results, and Games-Howell post hoc results for the four variables.

The open-ended question from Section D gave us more insight into what influenced our managers' decision. We found that the themes identified in the four conditions with visible content (FD, FO, MD, and MO) were different from those identified in the two control conditions which had no posts. We calculated the percentage of managers in each condition who mentioned responses falling under each theme. We include the themes having at least 10% in one or more study conditions. Tables 5.4 and 5.5 illustrate the main themes and the percentages of managers per condition with responses in each theme. The cells in each table are colour-coded to show where the most positive and the most negative impressions existed.

Table 5.6: **RQ2:** Effect of of the *representation type* × *gender of candidate* interaction on the hiring decision (V1-HIR) and 3 compound variables. The table shows two-way ANOVA test results. F: f-distribution, p: significance value, Partial η^2: effect size.

V#	***Two-way ANOVA***		
	F(2, 354)	p	Partial η^2
V1-HIR	0.722	0.486	0.004
V2-PER	0.148	0.863	0.001
V3-POS	0.324	0.724	0.002
V4-NEG	0.219	0.804	0.001

For the conditions with visible content, we extracted a total of 14 themes; eight themes met the 10% threshold. We found that managers inferred positive characteristics about the candidate: having positive personalities, having a positive lifestyle, having matured over time, having an intellect which indicated that they have the required qualifications for a job, and having no 'red flags' on their profile. These positive characterizations were more prevalent in the conditions with *decayed* representations. On the other hand, managers also inferred negative qualities about the candidate. They noted unprofessionalism, concerns about participation in protests, and disapproval of their partying lifestyle. These negative characterizations were more prevalent in the conditions with *original* representations.

We identified eight themes in the control conditions with no visible posts, FC and MC; six themes met the 10% threshold. Nearly half of our managers reported that they made their hiring decision based solely on details from our given scenario (i.e., that the candidate possessed knowledge and qualifications, and passed the interview). A quarter of managers viewed the fact that the candidate's profile is private as a positive trait, while another quarter thought there was not enough information available to make a judgment. Some managers mentioned that there was nothing negative on the profile, while others made assumptions about the user and their past history (e.g., the candidate does not use the account, moved to another platform, had problematic content in the past, or had deleted all of their content). A small percentage questioned the lack of content; they thought it was as worrisome as having inappropriate content and they would have appreciated more transparency.

In summary, we saw more negative comments in the original conditions. In the control conditions, managers ignored social media or made assumptions, and in the decayed conditions, they were more likely to have a positive impression.

5.4.2 RQ2 Analysis and Results

RQ2: *How does the gender of the candidate influence managers' hiring decisions?*

To answer our second research question, we examined the interaction effect of our two independent variables (*representation type* and *gender of candidate*) on V1-HIR to V4-NEG using two-way ANOVA. We found no significant interactions between the two independent variables on any of the four dependent variables. Table 5.6 shows two-way ANOVA results on the four variables.

5.4.3 RQ3 Analysis and Results

RQ3: *How do managers' demographics such as gender and age impact their hiring decisions?*

To answer our third research question, we looked for interactions between the *age* or *gender* of the managers and our *two independent variables* on our four dependent variables. We ran two-way ANOVA to look for an interaction effect on V1-HIR, V2-PER, V3-POS, and V4-NEG. We categorized age into three classes. We had a minimum age of 23 and a maximum of 64. Given this range and based on the distribution, we coded the variable into the following three classes: 23-36 years old (class 1 with n= 155), 37-50 (class 2, n= 153), and 51-64 (class 3, n= 52). Instead of coding the variable based on equal age ranges (i.e., 13-year blocks: 23-36, 37-50, 51-64), we could have coded based on equal numbers of participants per block, but we would have had very skewed ranges. Our choice seemed to more accurately reflect generational differences, and seemed reasonable given that the smallest group had 52 participants.

Results of two-way ANOVA explained below are illustrated in Table 5.7.

1. ***Representation type*** $\times$ **Gender of manager**: We found no significant interactions between these two variables on any of the V1-HIR – V4-NEG responses.

2. ***Representation type*** $\times$ **Age of manager**: We found no significant interactions between these two variables on any of the V1-HIR – V4-NEG responses.

3. ***Gender of candidate*** $\times$ **Gender of manager**: We found no significant interaction between these two variables on any of the V1-HIR – V4-NEG responses.

Table 5.7: **RQ3:** Effect of four interactions on the four dependent variables. The table shows two-way ANOVA test results for each.

Interaction	***V#***	***Two-way ANOVA***		
		F-distribution	p	Partial η^2
Representation Type	**V1-HIR**	F(2, 353) = 1.171	0.311	0.007
	V2-PER	F(2, 353) = 0.301	0.740	0.002
$\times$	**V3-POS**	F(2, 353) = 0.759	0.469	0.004
Manager Gender	**V4-NEG**	F(2, 353) = 0.080	0.923	0.000
Representation Type	**V1-HIR**	F(4, 351) = 1.344	0.253	0.015
	V2-PER	F(4, 351) = 0.316	0.867	0.004
$\times$	**V3-POS**	F(4, 351) = 0.830	0.507	0.009
Manager Age	**V4-NEG**	F(4, 351) = 0.742	0.564	0.008
Candidate Gender	**V1-HIR**	F(1, 355) = 0.019	0.892	0.000
	V2-PER	F(1, 355) = 0.143	0.706	0.000
$\times$	**V3-POS**	F(1, 355) = 0.218	0.641	0.001
Manager Gender	**V4-NEG**	F(1, 355) = 0.049	0.825	0.000
Candidate Gender	**V1-HIR**	F(2, 354) = 1.911	0.149	0.011
	V2-PER	F(2, 354) = 1.658	0.192	0.009
$\times$	**V3-POS**	F(2, 354) = 2.236	0.108	0.012
Manager Age	**V4-NEG**	F(2, 354) = 5.387	0.005	0.030

4. ***Gender of candidate*** $\times$ **Age of manager**: We found no significant interaction between these two variables on V1-HIR, V2-PER, and V3-POS responses.

   We found a statistically significant interaction between gender of the candidate and age of the managers on V4-NEG, F(2, 354) = 5.387, p = 0.005, partial η^2 = 0.030. Figure 5.6 shows the interaction between these two variables for V4-NEG (negative aspects of online reputation). As follow-up, an analysis of simple main effects for age was performed. Managers in the middle age group were more accepting of the female candidate (i.e., they minimized her negative online activities) ($p = 0.015$), while managers in the older group were more accepting of the male candidate ($p = 0.040$).

5.4.4 RQ4 Analysis and Results

RQ4: *How do OSNs and online reputation influence managers' real-life hiring decisions?*

To answer our fourth research question, we explore managers' reported use of OSNs and online reputation in their real-life practices (Section G of the questionnaire).

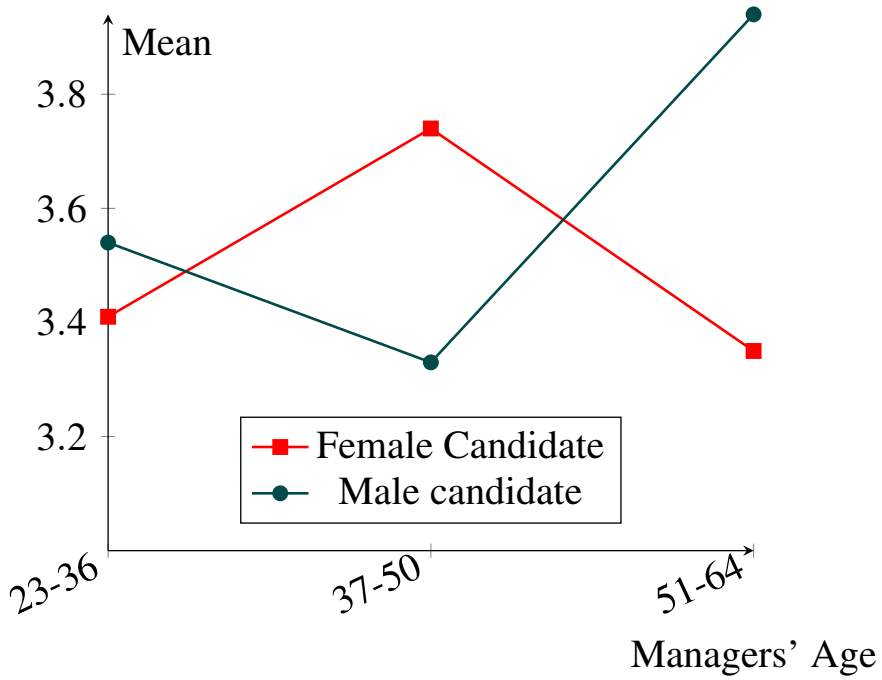


Figure 5.6: RQ3: Interaction effect of managers' age and candidates' gender on V4-NEG.

Table 5.8: **RQ4:** Number of managers who check each OSN platform during real-life hiring decision.

OSN Platform	*FD (n=68)*	*FO (n=61)*	*MD (n=61)*	*MO (n=62)*	*FC (n=55)*	*MC (n=53)*	*Total (n=360)*	*Total in %*
Facebook	60	51	53	56	48	44	312	87%
LinkedIn	38	33	37	34	32	34	208	58%
Instagram	30	23	27	29	21	25	145	40%
Twitter	31	29	22	23	21	24	150	42%
Google+	13	5	5	8	9	6	46	13%
Snapchat	6	5	5	4	4	0	24	7%
Tumblr	4	2	2	1	2	0	11	3%

Looking up candidates online. Most of our managers look up job candidates online before hiring them, as shown in Figure 5.7. They also check the various social media profiles for job candidates, as summarized in Table 5.8. The number of managers checking Facebook was particularly high. LinkedIn came a distant second, followed by Instagram and Twitter.

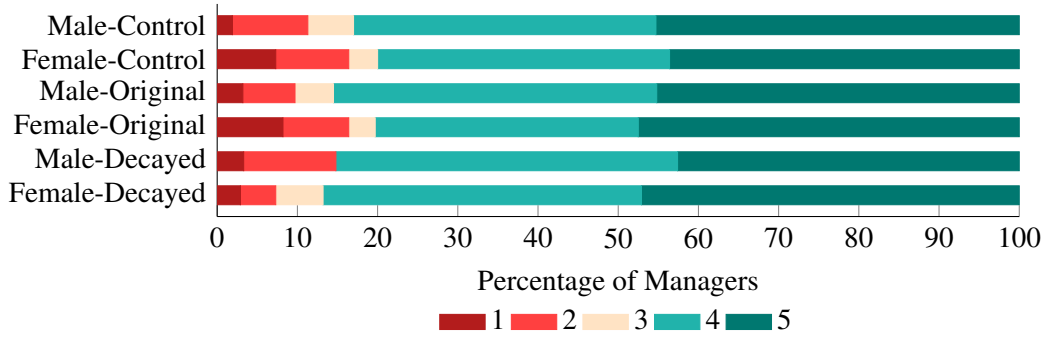


Figure 5.7: **RQ4:** Likelihood of managers looking up candidates online in real life (1 = Extremely unlikely, 5 = Extremely likely).

Table 5.9: **RQ4:** Main themes of how online reputation influences managers' real-life hiring decisions.

	To look for red flags	***Reflects personality***	***To see how they would fit in work***	***General influence***	***No influence***
FD (n= 68)	38%	35%	15%	12%	16%
FO (n= 61)	44%	46%	0%	26%	26%
FC (n= 55)	55%	29%	18%	13%	20%
MD (n= 61)	34%	38%	18%	16%	16%
MO (n= 62)	47%	37%	18%	15%	19%
MC (n= 53)	66%	25%	21%	19%	13%

How online reputation influences hiring decisions. We identified a total of eleven themes in responses to our open-ended question about how managers use social media in hiring decisions. Six themes were excluded, since they were below the 10% threshold. Table 5.9 shows the remaining five themes. Most managers used social media to look for 'red flags' that would sway their hiring decision. Many thought that social media would enable them to know the person better as it reflects personality and lifestyle. Others thought it would allow them to see how the candidate would fit in the job and with the company's values. Some reported that it plays a role in general, while relatively small minority thought that social media played no role at all in the hiring process.

Table 5.10: **RQ4:** Likert-scale responses for the likelihood that managers hire a candidate who posts specific types of content in real life (n= 360, 1 = Extremely unlikely, 5 = Extremely likely). Md: median, μ: mean, and SD: standard deviation

Online content type	***Md***	μ	***SD***
Could damage company's reputation	1.0	1.7	(1.0)
Against company's values	1.0	1.8	(1.1)
Shows participation in illegal activities	1.0	1.6	(1.1)
Shows inappropriate content	2.0	2.0	(1.1)
Shows controversial content	2.0	2.4	(1.0)
Shows negative comments about past employers	2.0	2.1	(1.0)

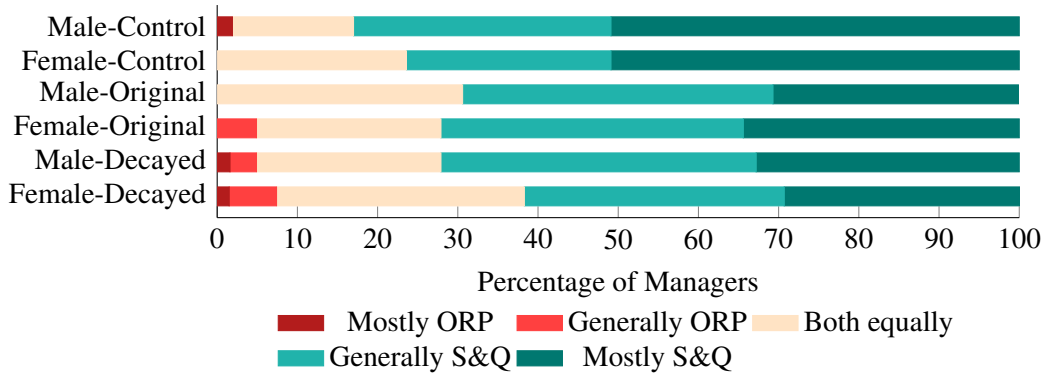


Figure 5.8: **RQ4:** Skills and qualifications (S&Q) vs. online reputation (ORP) in the hiring process in real life.

Skills & qualifications vs. online reputation. Figure 5.8 shows the most important factor identified by managers when making a hiring decision. Managers reported a candidate's skills and qualifications are more important than the candidate's online reputation. However, managers also said they were unlikely to hire a candidate who posts negative content (see Table 5.10).

Timeframe. 17% of our managers reported that all past online activities matter in terms of online reputation. 29% selected content from "1 year ago or less", 45% chose content posted "1-5 years ago", and only 1% said that content from "5-10 years ago" matters. Finally, 8% thought that "none" of a candidate's past online activities matter.

5.5 Discussion

Our results confirmed and also contradicted previous research. Addressing our research questions, we found:

RQ1: Decay representations (that apply gradual shrinking) positively impacted impression and hiring decisions, except for aspects of online reputation that could be perceived as negative where the Control representation led to a more positive impression.

RQ2: The candidate's gender had no significant effect on hiring decisions nor evaluation of the candidate's traits or online reputation.

RQ3: The managers' gender had no effect, but the managers' age had some limited impact on how male and female candidates were perceived.

RQ4: OSN content has considerable influence on real-life hiring practices.

We present implications of our results and suggestions for ORM and usable privacy communities.

5.5.1 Online reputation shapes hiring decisions

Consistent with previous research [28, 63, 73, 89, 125, 133], we found that online reputation continues to inform hiring decisions today. Interestingly, we saw a disconnect between managers' reported priorities and their actual assessment.

In Section 5.4.4, managers indicated that skills and qualifications of a candidate are more important than online reputation. However, in Section 5.4.1, we saw that their hiring decision significantly differed based on whether the profile hid some or all negative content from view compared to having the same content fully visible. Moreover, their responses in Table 5.9 and Table 5.10 suggest that online activities are very likely to influence their judgment.

Our managers reported that only more recent online activities (within 5 years) are relevant in their hiring decision. However, the significant difference between hiring decisions for the original and decayed profile indicates that posts dating 11 years ago were still relevant in their assessment. It is plausible that they did not carefully read the dates of the posts on the prototype. This may also happen in real life though, where managers could be unduly influenced by older posts without even realizing it. Decay representations could

help address this problem. We note that it is unclear whether managers' chosen timeframe for relevance of content in Section 5.4.4 is due to their actual preferences or due to the fact that it is fairly difficult and time-consuming to scroll through profiles to access older content. Given our data, we cannot confirm which reason contributed to their responses.

Furthermore, 84% of managers would look up their candidate online before hiring them. Managers are more likely to look up the candidate's Facebook or "social" profile instead of his/her LinkedIn or "professional" profile. While Facebook was previously reported as popular in the literature [34], it is interesting to see how much more popular it is than professional platforms such as LinkedIn.

This suggests that employers are interested in a prospective employee's personal life as well as professional side. Managers might think that looking at a LinkedIn profile is similar to looking at a candidate's résumé, but they are interested in knowing what is not articulated in a résumé. This aligns with the argument that employers think that checking OSNs provides them with answers to questions that are inappropriate to ask in an interview [133].

Job candidates might actively ensure that their professional profile is presentable for prospective managers, but managers are instead looking at profiles that candidates are less likely to actively manage [101, 147, 152]. The question remains, how can we eliminate negative consequences that might result from unintended data availability on OSN profiles? How can we empower users with tools that require their minimal engagement [145]?

5.5.2 Influence of Demographics

Candidates' demographics: Based on previous research [6, 66], we suspected there may be an effect of candidate gender on responses, but we were surprised to see that it did not play a role in our managers' hiring decisions, as shown in Section 5.4.2. Reflecting on this discrepancy, we note that we did not explicitly specify the gender of the candidates on the profile information or in our study scenario. Therefore, gender inference was open to participants and was based on their own interpretations of the profile. In addition, we controlled profile content so that only gender itself was varied. Managers saw the same content attributed to either a male or female candidate rather than different types of content that may show gendered behaviors [139]. The types of content posted by each gender may more strongly influence attitudes than gender itself. We did, however, see some interaction

between candidate gender and manager age (discussed below), suggesting that this issue is nuanced.

We note that for the purpose of this study, we focused on varying gender of the candidate rather than other demographics. We intentionally avoided including other characteristics known to cause bias (e.g., candidates with different ethnicity type [43, 44], younger/older candidate [40, 43]) to limit confounding factors, but these could also be explored.

Managers' demographics: The managers' demographics explored had limited impact on hiring decisions. Middle-aged managers minimized negative online activities of the *female* candidate, while older managers minimized negative online reputation of the *male* candidate, suggesting that stereotypes still exist and can affect the hiring process. Although the literature [40, 43, 44] suggests bias when there are differences between gender, race, or age of managers and candidates, it was interesting to see an interaction between two different characteristics (gender and age) in our study. Other factors such as ethnicity, cultural background, or work domain might also be worth investigating.

5.5.3 Decay representations improved impression

We found that the decaying and control representations led to significantly more positive hiring decisions than the original representation. Managers also had an overall more positive impression of both personality traits and the online reputation of the candidate with a decayed profile. Aligning with the literature [33, 34, 69, 73], our managers inferred characteristics related to the candidate's personality and lifestyle from social media, and their interpretations informed their hiring decision.

Although hiding posts appears to be the best option when online reputation contains negative content, it gave room for worrisome and potentially wrong assumptions about the online history of the candidate. Decayed profiles also performed reasonably well (Table 5.3) and have a clear advantage for the other two compound variables.

Managers were interested in seeing more about the person than what is available from an in-person interview [133]. Additionally, a perceived lack of information could sway their decision against the candidate or lead to misjudgment.

Overall, compared to hiding or showing the original content, managers who saw the decayed content gave the candidates the benefit of the doubt when they saw the other positive dimensions related to them.

5.6 Limitations

Crowdsourcing data is generally accepted within usable security and HCI communities, however, biases may exist. For example, our manager population was comfortable using online platforms and this might have skewed results. Thus, results are not necessarily representative of the whole population.

In addition, our fictitious candidate's profiles were available to our managers as videos instead of interactive profiles. Video scrolling is different than having managers navigate the profile themselves (e.g., speed of scrolling, what they pay attention to). However, creating an interactive UI was not feasible for the following reasons: 1) we could not enforce that everyone saw the same content, 2) we could not use an actual FB profile because we could not back-date posts by several years. We also needed to implement the shrinking and not allow users in those conditions to access the full resolution posts; this was not possible using the FB API, 3) we could have reproduced a fake interface but ran into issues of how much needed to be functional (how broadly and how deeply) because we did not want users to run into broken links. In the end, we opted for a video, despite its apparent drawbacks, because it seemed the most reasonable compromise.

Furthermore, the use of fictitious candidates might have been disconnected from a real employment context for our managers. In our study scenario, we asked managers to assume that they have already interviewed the candidate and are deciding whether to hire them. Results might differ in real-life employment scenarios (where managers interview and form a connection with their candidate). Additionally, in assuming that the candidate was qualified, the scenario may have introduced a bias by making managers look for an excuse to "not hire" the candidate instead of looking for something that would sway them to "hire" the candidate. However, this does appear to be how most employers use social media checks [32, 87]. We also do not ask about every possible factor impacting the hiring process. We limited our factors to prevent participant fatigue. Moreover, in study the scenario, no specific job opening was given. We avoided restricting managers to a specific

position and job criteria because they come from different domains. Nevertheless, it would be worth exploring whether the nature of the job causes managers to be more critical with respect to social media content.

We plan to extend this work using different methodologies to collect more detailed responses and explore managers' interpretations of the decay representation.

5.7 Conclusion

We explored how *decay representations* and job candidate gender influence hiring decisions in the digital age. Decay representations apply within OSNs by gradually decaying older posts. In this study, we applied a decay representation that gradually *shrunk* posts as they became older. We created six prototypes of fictitious job candidate social media profiles, varying the gender of the profile owner and the representation of the profile. We conducted a 3×2 between-subjects study with 360 managers in the US. We found no significant effect of candidate's gender on managers' decisions or perceptions. We also found a limited impact of managers' gender and age on the results. On the other hand, we found that decaying or hiding posts resulted in significantly more positive hiring decisions and assessment than the original representation. However, the *decay representation* also further promoted a positive impression of candidates while minimizing questions relating to lack of content. The *decay representation* helped managers overlook questionable past online activities of our fictitious candidates compared to the current format of social media posts, suggesting that this type of UI representation can be important to users when maintaining online reputation. We further found that our managers check OSNs in their real-life practices, emphasizing the need for tools that help users easily dissociate from past online content.

Chapter 6

The Influence of Decay Representations on Hiring Decisions - Interviews

In this chapter, we extend the work presented in Chapter 5. Through triangulation, we attempt to validate and deepen our understanding of the results from the previous study. Thus, the study design is closely related to the study in Chapter 5, except that semi-structured interviews are the main source for data collection. The new analysis includes content analysis of qualitative data, and the overall findings are compared with those from the previous chapter.

6.1 Research Questions

We address the same research questions from the previous chapter using different methods. We believe that the difference in the study design and analysis will provide more comprehensive answers to our research questions.

The research questions are:

RQ1: How does decaying a candidate's OSN profile influence managers' hiring decisions?

RQ2: How does the gender of the candidate influence managers' hiring decisions?

RQ3: How do managers' demographics such as gender and age impact their hiring decisions?

RQ4: How do OSNs and online reputation influence managers' real-life hiring decisions?

6.2 Methodology

The study was cleared by our Research Ethics Board (Clearance ID: 109256). We interviewed 48 people in Canada responsible for hiring decisions, primarily managers, in a 3×2 between-subjects study design (six conditions, 8 participants per condition). Participants were sequentially assigned to one of the six conditions by ensuring that we even out the

number of male/female participants viewing each profile. In our study scenario, we asked participants to assume that they have already interviewed the candidate and were deciding whether to hire them.

The study was pilot tested with four participants. One pilot session was run with a team manager who regularly performs recruitment duties as part of his job position. The other pilots were with graduate students who have usable security and privacy expertise. The pilot sessions helped us refine the progression of the session, the interview guide, and the post-study questionnaire[1].

6.2.1 Procedure

We collected participants' feedback verbally and through online questionnaires in a 60-minutes audio-recorded session. Prototypes and videos associated with each condition from the previous study were used: Female-decayed (FD), Female-Original (FO), Female Control (FC), Male-Decayed (MD), Male-Original (MO), and Male Control (MC). The decayed prototypes showed a Facebook profile with posts gradually and linearly shrinking in size. The original prototypes showed a Facebook profile with all posts displayed full size. The control prototypes showed a Facebook profile with no posts displayed on the timeline. The six prototypes were presented to participants as videos. A detailed explanation of the prototypes and content of the profiles were presented in Section 5.2.1 of the previous chapter. We collected the same demographics information (from Section 5.2.2-A) and collected more data related to participants' experience in general. We presented the same survey (detailed in Section 5.2.2) as a post-study questionnaire. In addition, we wanted to collect as much feedback as possible from our total sample size related to the decay representations. Therefore, participants not assigned to the FD and MD conditions were shown the videos associated with those two conditions at the very end of the interviews, to avoid any priming effects. To clarify, a session unfolded as follows:

1. The researcher introduced the purpose of the study.

2. Participants completed the consent form.

[1]Any changes to the post-study questionnaire were also applied to the questionnaire in the previous study in Chapter 5 as pilot sessions of both studies ran concurrently

3. Participants filled out a demographic questionnaire (Section A of Appendix C).

4. Participants were verbally asked some questions related to their job position and experience.

5. Participants were verbally introduced to the study scenario and watched the video displaying the prototype for their study condition. The scenario assumed that participants had already interviewed the candidate and found that he/she was qualified, and now they are looking at his/her Facebook profile and deciding whether to hire him/her.

6. Participants interacted with the video and reviewed it as much as they wanted, then made a hiring decision.

7. Participants were verbally asked follow-up questions to explore their hiring decision related to the applicant's profile shown in the video.

8. Participants were asked another set of questions exploring their real-life hiring practices in relation to online reputation.

9. Participants then completed a post-study questionnaire (Sections B, D, E, F, and G of Appendix C).

10. Participants not assigned to the FD and MD conditions were shown the videos associated with those two conditions. Participants were then asked few questions related to their interpretation of the representation and how it would influence their hiring decision. The purpose of the representation was explained afterwards and participants were asked for comments and feedback.

6.2.2 Interview Questions

Interview questions investigated participants' hiring decisions related to our fictitious applicants. We explored what influenced their decisions and what they thought of the online activities of the applicants. Questions further probed how online reputation shapes their

decisions in real life. For example, we examined how they weigh the applicant's qualifications against online reputation and which aspects of online reputation are influential in their assessment and hiring decisions. The full interview guide is available in Appendix D.

6.2.3 Recruitment and Participants

We recruited participants with specific demographics. To qualify, participants must: (1) be in managerial positions or responsible for hiring decisions, (2) have experience using social media platforms, (3) be able to speak/read/write English fluently.

Recruitment was done in Canada through a social media page for advertising Carleton university's HCI studies. We also advertised the study poster on Facebook and LinkedIn on groups that might have potential participants, after approval from group moderators. The researchers asked organizations with which they have an existing relationship if they are willing to send the recruitment email to their members or industry partners.

Participants were compensated $20 in cash for their time. Before beginning the session, participants read and signed a consent form explaining the purpose and the procedure of the study and reminding them that the session will be audio-recorded. Personally identifiable information collected from participants were limited to their voice; responses were pseudo-anonymized and non-attributable.

For analysis and reporting results, participants are assigned a username that is not linked to their identity. Usernames are generated according to participants' assigned prototype (e.g., Female-Decayed: FD1–FD8, Female-Original: FO1–FO8, Female-Control: FC1–FC8, Male-Decayed: MD1–MD8, Male-Original: MO1–MO8, Male-Control: MC1–MC8).

All participants had at least an undergraduate or college degree. Most participants had a post-graduate degree. They were also mostly "very familiar" or "moderately familiar" with social media. They came from 19 various fields, but most were in government, education, technology & computers, followed by engineering, business, the service sector, followed by other miscellaneous fields. Table 6.1 summarizes participants' demographics.

Table 6.1: Participants' demographics per condition.

	Category	*FD*	*FO*	*MD*	*MO*	*FC*	*MC*
Gender	Male	3	4	4	5	5	4
	Female	5	4	4	3	3	4
	Total	8	8	8	8	8	8
Age	Minimum	19	22	21	19	21	19
	Maximum	62	51	57	72	63	44
	Mean	40	39	40	44	41	32
	Std Dev.	14	9	13	18	12	10
Average Number of hiring decisions per year	Minimum	1	3	1	1	2	1
	Maximum	12	15	4	5	25	25
	Mean	6	9	2	3	9	6
	Std Dev.	5	5	1	2	6	8
Years of work experience	Minimum	1	2	0.5	0.5	0.5	0.5
	Maximum	15	22	11	28	40	20
	Mean	8	8	5	12	10	5
	Std Dev.	5	7	4	9	13	7

6.3 Analysis Plan

To answer our RQ1 and RQ4, we ran a statistical analysis and conducted qualitative analysis. For our RQ2 and RQ3, we ran only a statistical analysis to look for main effects. We include below a description of our two analyses.

6.3.1 Statistical Analysis

We ran the same statistical tests described in Section 5.3 on the four dependent variables that were based on 21 questionnaire questions: the *hiring decision* V1-HIR (ordinal data), and three compound scale data variables that evaluated *personality traits* of the candidate (V2-PER), *positive aspects* of online reputation (V3-POS), and *negative aspects* of online reputation (V4-NEG). A detailed description of the four variables and the calculation process for the three compound variables was described in Section 5.3 of Chapter 5. Table 6.2 shows descriptive statistics for the hiring question and for the 20 questions constituting the three compound variables.

Table 6.2: Likert-scale responses for the hiring decision and the 20 individual questions used in the compound variables (V2-PER, V3-POS, V4-NEG), organized by either *representation* or *gender*, with the highest means in bold red.

Q#	*Representation*									*Gender*					
	Decayed (n= 16)			*Original (n= 16)*			*Control (n= 16)*			*Female (n= 24)*			*Male (n= 24)*		
	Md	μ	SD	Md	μ	SD	Md	μ	SD	Md	μ	SD	Md	μ	SD
D1: Hiring	4.0	4.0	(1.0)	4.5	**4.2**	(0.9)	3.5	3.8	(0.9)	4.0	**4.2**	(0.8)	4.0	3.9	(1.1)
E1: Hard working	3.0	**3.3**	(0.8)	3.0	2.8	(0.8)	3.0	2.9	(0.5)	3.0	3.0	(0.8)	3.0	3.0	(0.7)
E2: Self-motivated	4.0	**3.9**	(0.8)	4.0	3.6	(0.9)	3.0	3.0	(0.6)	3.0	**3.5**	(1.0)	3.0	3.4	(0.8)
E3: Loyal	3.0	**3.5**	(0.7)	3.0	3.1	(0.8)	3.0	2.9	(0.6)	3.0	3.2	(0.9)	3.0	3.2	(0.6)
E4: Dependable	3.0	**3.3**	(0.9)	3.0	2.8	(1.0)	3.0	3.0	(0.6)	3.0	3.0	(1.0)	3.0	3.0	(0.8)
E5: Team-oriented	4.0	**3.6**	(0.7)	3.0	3.4	(1.0)	3.0	2.9	(0.6)	3.0	3.3	(0.9)	3.0	3.3	(0.8)
E6: Confident	4.0	4.0	(0.6)	4.0	**4.1**	(0.6)	3.0	3.1	(0.9)	4.0	**3.8**	(1.0)	4.0	3.6	(0.6)
E7: Adept communicator	3.4	3.0	(0.6)	4.0	**3.7**	(0.6)	3.0	2.9	(0.6)	3.0	3.3	(0.8)	3.0	**3.4**	(0.6)
E8: Respectful to work ethics	3.0	2.9	(0.6)	2.5	2.6	(0.8)	3.0	2.9	(0.5)	3.0	2.7	(0.7)	3.0	**2.8**	(0.6)
E9: Flexible	3.0	**3.4**	(0.6)	3.0	3.3	(0.7)	3.0	3.0	(0.6)	3.0	3.1	(0.7)	3.0	**3.4**	(0.6)
E10: Cautious	3.0	2.6	(0.9)	3.0	2.5	(0.6)	3.0	**3.5**	(0.7)	3.0	**3.0**	(1.0)	3.0	2.8	(0.8)
E11: Positive lifestyle	5.0	**4.6**	(0.6)	4.0	4.2	(0.5)	3.0	3.3	(0.9)	4.0	3.9	(1.0)	4.0	**4.1**	(0.7)
E12: Positive attitude	4.5	**4.4**	(0.7)	4.0	4.1	(0.6)	3.0	3.2	(0.9)	4.0	3.8	(1.0)	4.0	**4.0**	(0.7)
E13: Professionalism	3.0	2.9	(0.8)	3.0	2.6	(1.0)	3.0	**3.1**	(0.8)	3.0	2.9	(0.9)	3.0	2.9	(0.9)
E14: Volunteering	3.0	**3.1**	(1.1)	3.0	2.9	(0.6)	3.0	2.8	(0.7)	3.0	2.6	(0.9)	3.0	**3.2**	(0.6)
E15: Openness	4.0	4.1	(0.9)	4.0	**4.2**	(0.7)	3.0	3.1	(0.8)	4.0	3.8	(1.0)	4.0	3.8	(0.8)
E16: History of drugs	3.5	3.5	(0.7)	4.5	**4.1**	(1.1)	3.0	3.4	(0.8)	3.0	**3.7**	(0.9)	3.5	3.6	(0.9)
E17: History of alcohol	3.0	3.2	(0.8)	2.0	1.9	(1.1)	3.0	**3.4**	(0.8)	3.0	2.8	(1.3)	3.0	**2.9**	(1.0)
E18: Offence to employer	2.0	2.8	(1.2)	2.0	1.9	(1.1)	3.0	**3.4**	(0.8)	2.0	2.5	(1.3)	3.0	**2.8**	(1.2)
E19: Inappropriate content	4.0	**3.9**	(1.0)	3.0	3.4	(1.1)	3.0	3.7	(0.9)	3.0	3.6	(1.2)	3.0	**3.8**	(0.9)
E20: Controversial content	3.0	3.1	(1.1)	3.0	2.8	(1.1)	3.5	**3.5**	(1.0)	3.0	**3.3**	(1.1)	3.0	3.0	(1.0)

6.3.2 Qualitative analysis

The average length of an interview session was 45 minutes. We had a total of 26 hours of audio files to transcribe. We selectively transcribed ten interviews manually using InqScribe [2], focusing our efforts on the relevant parts. The rest of the interviews were fully transcribed using Trint [3], an online transcription software, then manually edited for accuracy.

For this, we used a *deductive* process of content analysis [38] to compile and categorize excerpts using NVivo 12 Plus software [4]. In a deductive content analysis, themes emerge from the raw data through repeated examination. It allowed us to iteratively identify themes derived from raw data. Since we had already pre-identified the main categories we were looking for in our interview guide, our main goal was to identify subcategories and refine these into broader themes and understand the relationship between them. We clarify our process as follows:

[2] `https://www.inqscribe.com/`

[3] `https://www.trint.com/`

[4] `https://www.qsrinternational.com/nvivo/home`

1. We grouped the relevant interview questions into three main categories (i.e., assessment of the candidate, influence of representation, influence of online reputation).

2. We coded each excerpt falling into each main category.

3. From these codes, we identified subcategories (e.g., *give advice about sharing*, *displays low social media awareness*).

4. We then grouped the subcategories into themes (e.g, *Discuss social media use with applicant*) and identified relationships between these themes.

5. We repeated the previous steps until satisfied that it accurately represented our data.

6. We further refined our results by looking for key patterns in our broad themes and by identifying links between them.

The main researcher summarized responses, coded the data, and extracted main themes and key patterns across the themes. The research supervisor was involved in identifying the key patterns as well as in identifying links between themes to better interpret the data and handle any complicated or unclear cases.

We had a total of 784 excerpts which were condensed into 67 broad themes. The 67 themes were further condensed after identifying relationships and links between them. We clarify which questions of the interview guide were considered for content analysis and indicate the number of extracted themes for each of RQ1 and RQ4 in the following sections.

6.4 Results

We first present our participants' general preferences when it came to the 20 traits for Section B of the questionnaire (answered before watching the video). Figure 6.1 shows the 10 personality traits and their importance to our participants. Figure 6.2 similarly shows the importance of the 10 aspects of a candidate's online reputation. Overall, participants report being more interested in a candidate's personality than his/her online reputation and activities.

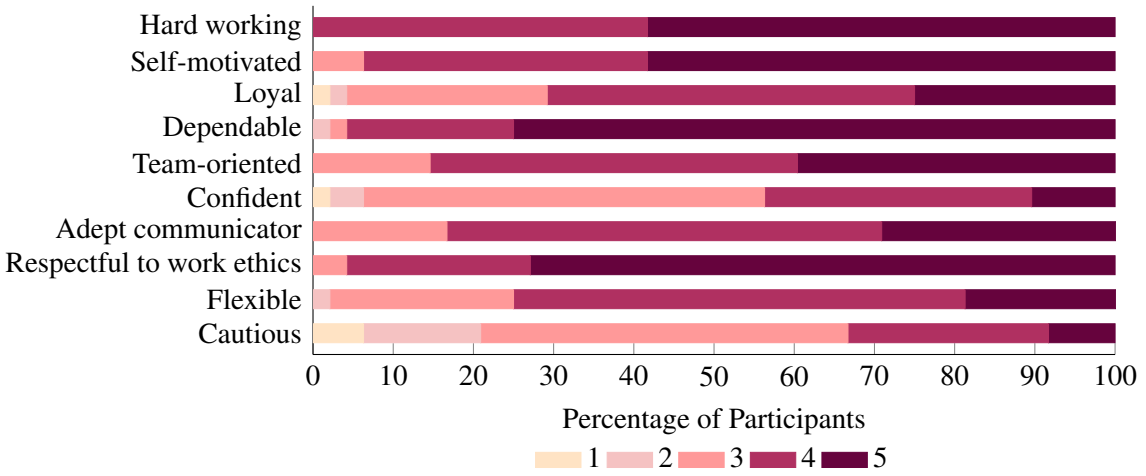


Figure 6.1: Ten personality traits as rated by participants (1 = Not at all important, 5 = Extremely important).

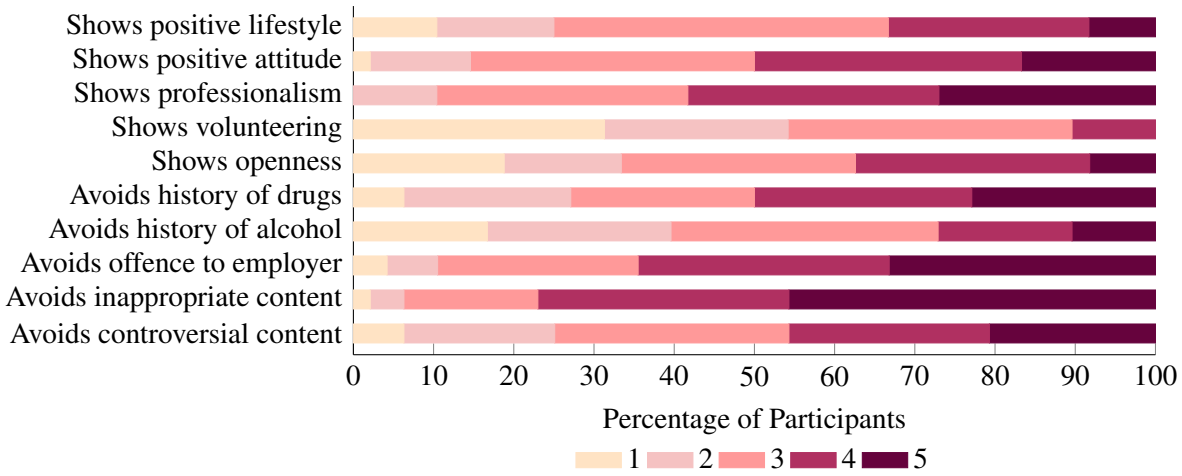


Figure 6.2: Ten aspects of online reputation as rated by participants (1 = Not at all important, 5 = Extremely important).

6.4.1 RQ1 Analysis and Results

RQ1: *How does decaying a candidate's OSN profile influence managers' hiring decisions?*

To answer our first research question, we looked at questions from both the questionnaire and the interviews that considered participants' hiring decision and their assessment of the candidate. We used both statistical analysis and content analysis to interpret our data. We first present results from the statistical analysis followed by results of the content analysis.

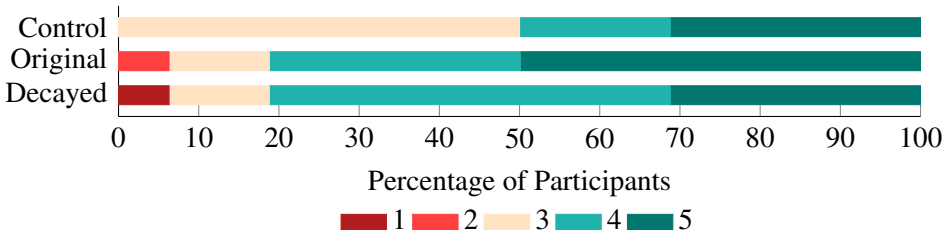


Figure 6.3: Likert-scale responses to the hiring decision question as provided by our participants per condition. 1= Very unlikely, 5= Very likely.

Statistical analysis:

We ran one-way ANOVA on V1-HIR to V4-NEG to look for main effects of our independent variable, representation type. The hiring decision made by participants in each of our three study conditions is shown in Figure 6.3.

Participants of both the decayed (81%) and the original (81%) representations were either somewhat or very likely to hire the candidate compared to those (50%) who saw the control representation. Interestingly, seeing an empty profile in the Control condition meant that half of participants were unsure about hiring the candidate, but none explicitly said they would not hire.

V1-HIR: The one-way ANOVA test[5] showed no significant effect ($p = 0.440$) of representation on V1-HIR. The descriptive statistics are as follows: Original ($\mu = 4.2$, standard deviation $(SD) = 0.9$), Decayed ($\mu = 4.0$, standard deviation $(SD) = 1.0$), and Control ($\mu = 3.8$, standard deviation $(SD) = 0.9$).

V2-PER: We found no significant effect of representation ($p = 0.063$) on participants' evaluation of the candidate's *personal traits*.

V3-POS, V4-NEG: We found a significant effect of representation on both V3-POS ($p = 0.003$) and V4-NEG ($p = 0.025$). We followed up with Games-Howell post-hoc analysis for pairwise comparisons. We found that participants who saw the Decayed representation (D) gave more positive responses when evaluating the candidate's *positive aspects of online reputation* compared to the Control (C) representation, a statistically significant

[5]We also ran independent samples Kruskal-Wallis test on V1-HIR. The KW test showed no significant effect ($p = 0.310$) of representation on V1-HIR.

Table 6.3: RQ1: Effect of representation type on each of the four dependent variables. The table shows mean and standard deviation (SD), one-way ANOVA test results, and Games-Howell pairwise results. The highest means are in bold red and significant results are highlighted in gray. F: f-distribution, p: significance value, Partial η^2: effect size, D: Decayed, C: Control, O: Original representation.

	Mean (SD)			***ANOVA***			***Pairwise Comparison***		
V#	D	C	O	F(2, 45)	p	Partial η^2	D - O	D - C	O - C
V1-HIR	**4.0 (1.0)**	3.8 (0.9)	4.2 (0.9)	0.837	0.440	0.036	0.754	0.850	0.383
V2-PER	**3.4 (0.5)**	3.0 (0.5)	3.2 (0.3)	2.946	0.063	0.116	0.314	0.086	0.559
V3-POS	**3.8 (0.5)**	3.1 (0.7)	3.6 (0.4)	6.660	0.003	0.228	0.371	0.010	0.066
V4-NEG	3.3 (0.6)	**3.5 (0.8)**	2.8 (0.7)	4.013	0.025	0.151	0.102	0.703	0.042

result ($p = 0.01$). However, there was no significant difference between the Decayed–Original (O) and the Control–Original pairs ($D > C, D \approx O, C \approx O$).

On the other hand, participants who saw the Control representation gave more positive responses when evaluating the candidate's *negative aspects of online reputation* (i.e., they minimized candidate's negative online reputation) compared to the Original representation, a statistically significant result ($p = 0.042$). However, there was no significant difference between the Decayed–Original and the Decayed–Control pairs ($C > O, D \approx O, D \approx C$). Table 6.3 shows means and standard deviation, ANOVA test results, and Games-Howell post hoc results for the four variables.

We compare results from this Chapter with those obtained from the crowdsourced study in Chapter 5 in the discussion section.

Content analysis:

We looked at participants' responses follow-up questions to understand what influenced their hiring decision and their impression of the candidate's personality and activities. We also looked at the questions that were relevant to the visual representation. We extracted a total of 43 themes for this set of questions. The relevant interview questions were[6]:

- Why would you hire or not hire the candidate?
- What influenced your hiring decision?

[6]These questions may have been rephrased slightly depending on the conversation flow with participants.

- With respect to personal qualities we talked about earlier, how does this person meet your criteria?
- What did you think of the candidate's online activities?
- How could you differentiate between older and most recent posts?
- Pointing out the shrinking effect: If you saw this representation, what does this mean to you? / what does it tell you?
- How would this visual representation influence your hiring decision?
- The intention of this is to put less attention on the older posts by making them smaller in size, what are your comments about this?

Hiring decision and assessment of the candidate: We categorized responses relating to their candidate assessment. We looked at how they perceived the candidate's online activities and their impression of the candidate's personality.

We observed similar responses from participants who saw the Decayed and the Original profiles. Their overall impression of the content was that it was normal, innocuous, and potentially concerning in the past but that the most recent content reflected maturity. However, they did notice what they described as *red flags* in both profiles. The red flags were mainly related to the negative employer post and it raised questions for them.

In the Decayed condition, participants thought that the red flags were mildly concerning but not a deterrent to getting the job; they would have a conversation with the candidate and judge from the person's attitude. On the other hand, participants from the Original condition mentioned that they would do further investigation and a more thorough reference check to decide on the candidate. In both conditions, many participants mentioned they would have a conversation with the candidate and advise them around privacy and clarify the boundaries between professional and personal life in their line of work.

In regards to personality, most participants in both conditions inferred many positive traits about the candidate, mainly around having a positive, loyal, well-rounded personality, and a few indicated that the candidate would be a good fit for their team. Some participants from the Original condition and a couple from the Decayed condition thought

Table 6.4: Main themes identified by participants' assessment of the candidate's personality and profile. Text in red refers to negative impressions or situations were questions were raised.

	Original	***Decayed***	***Control***
Red Flags	-Noticed -Triggered -Discuss with the applicant -Investigate further	-Noticed -Triggered, but not a deterrent -Discuss with the applicant -Judge in person	-No red flags -Investigate further
Content	-No influence -Concerns in the past/recent is mature -Nothing extreme or out of the ordinary	-No influence -Concerns in the past/recent is mature -Nothing extreme or out of the ordinary	-No influence -May be fake -Assumptions
Personality Impression	-Positive -Negative -Cannot tell	-Positive -Cannot tell	-Private -Cannot tell

it was hard to judge personality solely from a profile. A few participants from the Original condition inferred negative personality traits about the candidate, such as lack of respect, unprofessionalism, and irresponsibility; participants in the Decayed condition identified no such negative traits.

On the other hand, participants from the Control condition had very different impressions. The lack of posts on the Control profile did not signify a red flag for some participants and would not negatively influence their hiring decision. However, it did raise questions for others who would look up the candidate further and do more research about them. One group of participants appreciated the candidate's right to privacy, while another group started making assumptions about the candidate and the account. For example, they assumed the user either locked down their profile or hid all their posts, did not post at all, posted frequently but only to friends, or were not interested in social media. A few participants wondered if it was a fake profile or a made-up person.

Moreover, most participants in the Control condition found it hard to infer anything relating to the candidate or their personality, but a few thought the candidate is discreet and aware of their privacy.

Table 6.4 summarizes the overall observed themes per condition and participants' reaction in each one.

Influence of Representation Type: Participants from the Decayed condition had three main initial reactions when asked about the meaning of the visual representation. They either got its intention, thought it was a platform setting or a UI error, or paid no attention

to it. We asked them, hypothetically, what it would suggest to them if a new feature intentionally displayed profiles like this. They mostly said it would mean that the smallest posts are older or less relevant to the user. Some participants said that they wanted to enlarge or expand the smaller posts, oblivious to the privacy implications.

Most participants from the two other conditions who were shown the Decayed profile at the end of their session had a different reaction. They realized its intention and mentioned that it indicates age or relevance of the post.

After we explained the intention of the representation, there was two main opinions. Most participants expressed that it would be a good visual representation: a useful one that would remind the viewer that posts are getting older or becoming less relevant, or that it would make browsing more efficient if they want a glimpse of the user's activities. A few disliked the idea because they would ultimately want to see all posts in full size to not miss any red flags; these participants valued their desire for information more than the candidates' privacy.

In regards to its influence on their real-life hiring decisions, they mentioned that it would either have no influence on their practice of how they look up a candidate or that it would naturally lead them to put less emphasis on the candidate's past. A few mentioned that they would not want to miss anything.

RQ1 Results Summary

The Decayed and Original profiles were more helpful and let to more positive impressions of the candidate by participants compared to the Control profile. Between the Original and Decayed profiles, participants' impressions were similar except that participants in the Original condition identified a need to reference-check the candidate and investigate negative personality impressions. Participants in the Decayed condition were more likely to give the candidate the benefit of the doubt. On the other hand, the Control profile only suggested that the candidate is aware of privacy but participants could not infer any further impressions. It did not directly influence their assessment either positively or negatively but prompted them to make assumptions about the candidate or to look them up further.

6.4.2 RQ2 Analysis and Results

RQ2: *How does the gender of the candidate influence managers' hiring decisions?*

To answer our second research question, we examined the interaction effect of our two independent variables (*representation type* and *gender of candidate*) on our four dependent variables. We ran two-way ANOVA on V1-HIR to V4-NEG. We found no significant interactions between the two independent variables on any of the V1-HIR, V2-PER, and V3-POS variables.

We found a significant interaction between the two independent variables on V4-NEG (negative aspects of online reputation) which is illustrated in Figure 6.4. Participants who saw the Original profile of the Male candidate were significantly more accepting of the *negative* aspects of his online reputation than those who saw the Original profile of the Female candidate ($p = 0.015$).

Looking at the graph, other interesting patterns emerge. Notably that in the Decayed condition, we see no differences at all between genders, and in the Control condition, the female candidate is more highly rated for having no negative content in her profile. A larger sample might show statistical significance, but we are unable to verify at this stage.

Although we saw no significant effect when participants evaluated *positive* aspects of online reputation (V3-POS), we show the interaction in Figure 6.5 to compare with the observed patterns in Figure 6.4. Interestingly, similar patterns emerge; both candidates have close ratings in the Decayed condition and the female candidate is rated slightly lower in the Original condition. However, in the Control condition, the female candidate is rated lower for having no positive content in her profile.

Figure 6.6 shows the non-significant interaction for V2-PER. No obvious pattern emerges, other than the female candidate appears to be rated more highly than the male candidate in the Decayed condition. In each of these instances, it is unclear whether the pattern is a result of random chance or whether more data would confirm the pattern.

Table 6.5 shows the two-way ANOVA results on the four variables.

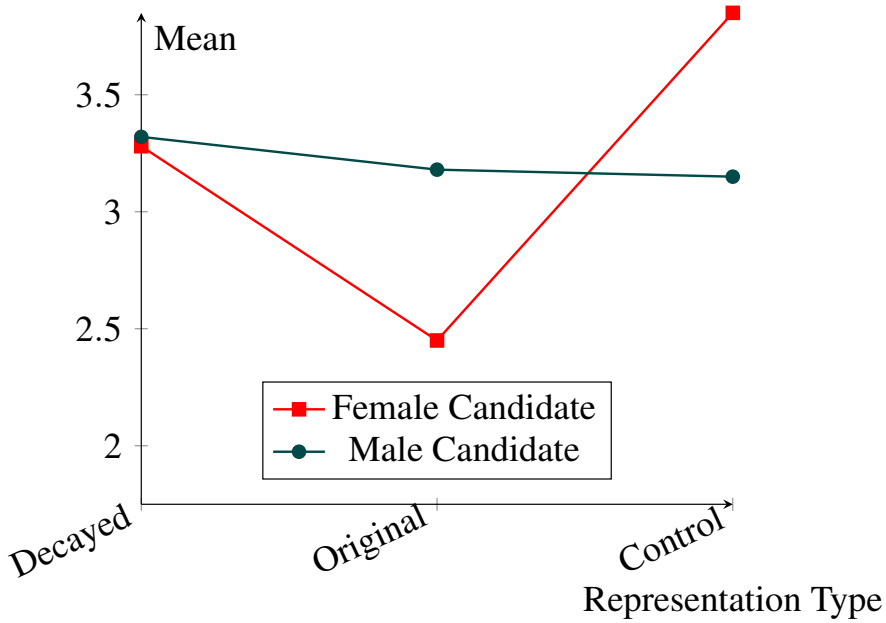


Figure 6.4: **RQ2:** Interaction effect of representation type and candidates' gender on *negative aspects* of online reputation (V4-NEG).

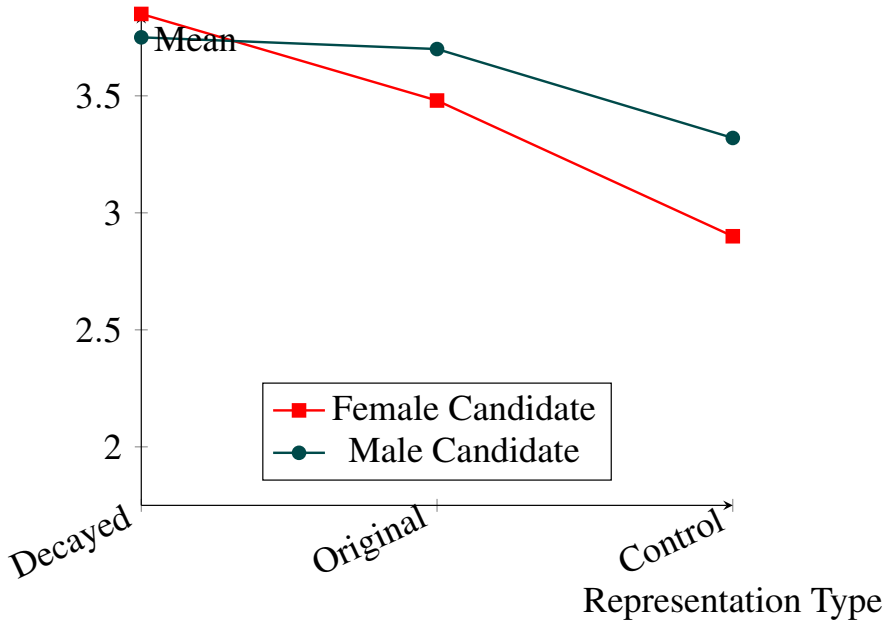


Figure 6.5: **RQ2:** Interaction of representation type and candidates' gender on *positive aspects* of online reputation (V3-POS).

Table 6.5: **RQ2:** Interaction effect of the *representation type* × *candidate's gender* on the four dependent variables. The table shows two-way ANOVA test results. Significant results are highlighted in gray. F: f-distribution, p: significance value, Partial η^2: effect size.

V#	***Two-way ANOVA***		
	F(2, 42)	p	Partial η^2
V1-HIR	0.285	0.753	0.013
V2-PER	0.998	0.377	0.045
V3-POS	0.949	0.395	0.043
V4-NEG	4.648	0.015	0.181

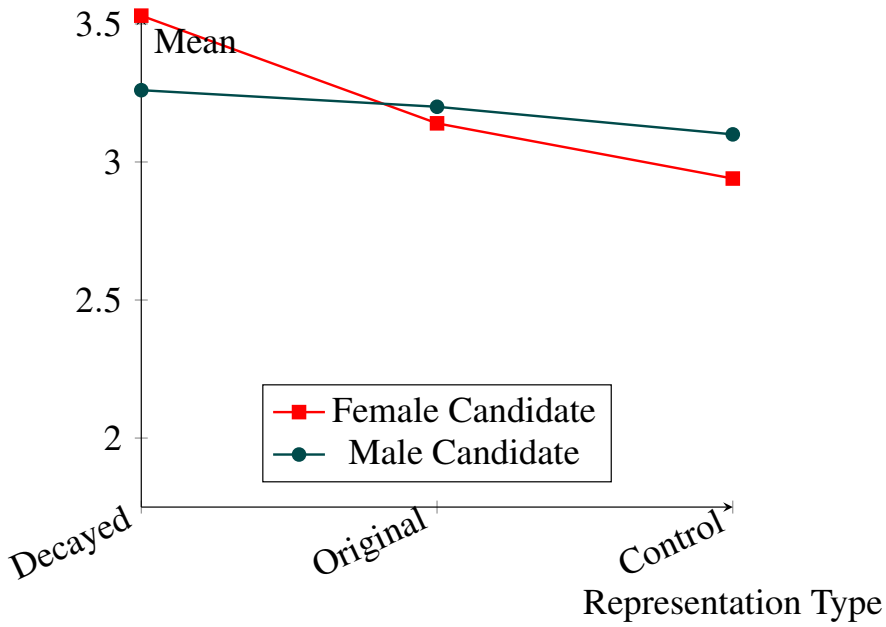


Figure 6.6: **RQ2:** Interaction of representation type and candidates' gender on *personality traits* (V2-PER).

6.4.3 RQ3 Analysis and Results

RQ3: *How do managers' demographics such as gender and age impact their hiring decisions?*

To answer our third research question, we looked for interactions between the *age* or *gender* of the participants and our *two independent variables* on our four dependent variables. We ran two-way ANOVA on V1-HIR to V4-NEG. We categorized age into three classes. We had a minimum age of 19 and a maximum of 62 plus one participant aged 72. Given this range and based on the distribution, we coded the variable into the following

Table 6.6: **RQ3:** Four interaction effects on the four dependent variables. The table shows two-way ANOVA test results for each.

Interaction	***V#***	***Two-way ANOVA***		
		F-distribution	p	Partial η^2
Representation Type	**V1-HIR**	F(2,42) = 0.492	0.615	0.023
	V2-PER	F(2,42) = 0.334	0.718	0.016
×	**V3-POS**	F(2,42) = 0.109	0.897	0.005
Participant Gender	**V4-NEG**	F(2,42) = 3.516	0.039	0.143
Representation Type	**V1-HIR**	F(4,39) = 0.278	0.890	0.028
	V2-PER	F(4,39) = 0.922	0.461	0.086
×	**V3-POS**	F(4,39) = 1.590	0.196	0.140
Participant Age	**V4-NEG**	F(4,39) = 0.522	0.720	0.051
Candidate Gender	**V1-HIR**	F(1,44) = 0.000	0.987	0.000
	V2-PER	F(1,44) = 0.009	0.926	0.000
×	**V3-POS**	F(1,44) = 0.334	0.566	0.008
Participant Gender	**V4-NEG**	F(1,44) = 0.017	0.896	0.000
Candidate Gender	**V1-HIR**	F(2,42) = 1.696	0.196	0.075
	V2-PER	F(2,42) = 0.826	0.445	0.038
×	**V3-POS**	F(2,42) = 0.633	0.536	0.029
Participant Age	**V4-NEG**	F(2,42) = 0.164	0.849	0.008

three classes: 19-33 years old (class 1 with n= 15), 34-48 (class 2, n= 21), and 49+ (class 3, n= 12).

Results of two-way ANOVA explained below are illustrated in Table 6.6.

1. ***Representation type*** × **Gender of participant**: We found no significant interaction between these two variables on V1-HIR, V2-PER, and V3-POS responses.

   We found a statistically significant interaction between representation type and gender of the participants on V4-NEG, F(2, 42) = 3.516, p = 0.039, partial $\eta^2 = 0.143$. Figure 6.7 shows the interaction between these two variables for V4-NEG (negative aspects of online reputation). As follow-up, an analysis of simple main effects for representation was performed. Female participants were significantly more accepting of the Control profiles when evaluating negative aspects of online reputation more than Male participants ($p = 0.025$).

2. ***Representation type*** × **Age of participant**: We found no significant interactions between these two variables on any of the V1-HIR – V4-NEG responses.

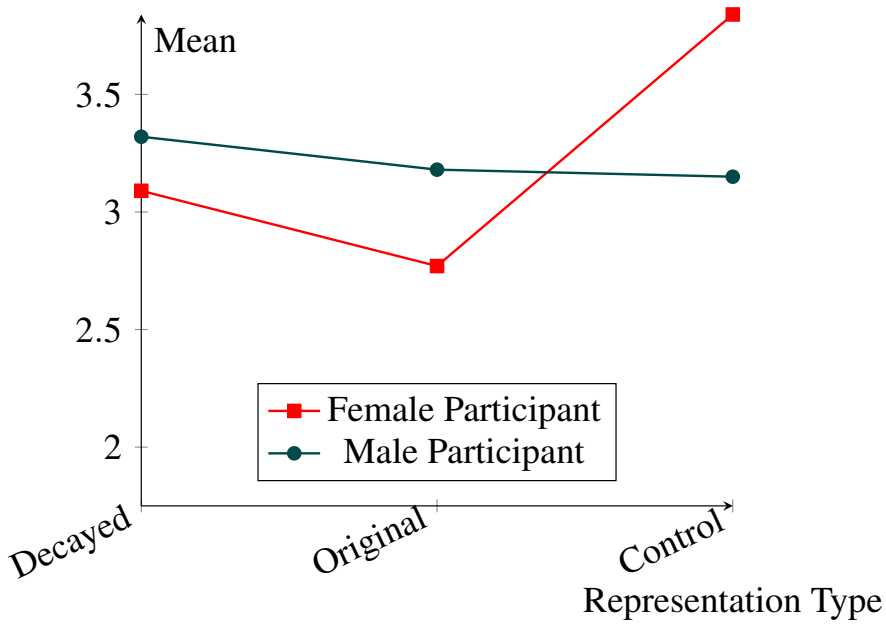


Figure 6.7: **RQ3:** Interaction effect of representation type and gender of participant on V4-NEG.

3. ***Gender of candidate*** × **Gender of participant**: We found no significant interaction between these two variables on any of the V1-HIR – V4-NEG responses.

4. ***Gender of candidate*** × **Age of participant**: We found no significant interaction between these two variables on any of the V1-HIR – V4-NEG responses.

6.4.4 RQ4 Analysis and Results

RQ4: *How do OSNs and online reputation influence managers' real-life hiring decisions?*

To answer our fourth research question, we explore participants' reported use of OSNs and online reputation in their real-life practices. These were explored using the questionnaire (Section G) where we report on descriptive statistics, and also through the interview questions where we report results from our content analysis.

Table 6.7: **RQ4:** Number of participants who check each OSN platform during real-life hiring decision.

OSN Platform	*FD*	*FO*	*MD*	*MO*	*FC*	*MC*	*Total*	*Total in %*
LinkedIn	7	6	4	6	7	6	36	75%
Facebook	5	4	5	3	2	3	22	46%
Twitter	4	3	3	2	4	2	18	38%
Instagram	2	2	1	-	1	3	9	19%
Google Search	2	-	1	2	1	-	6	12%
Google+	1	-	-	2	1	-	4	8%
Internal GC SM	-	-	-	-	2	-	2	4%
Personal websites	-	1	-	-	-	-	1	2%
Stack Overflow, Github	-	-	-	-	-	1	1	2%
None	-	1	-	-	-	2	3	6%

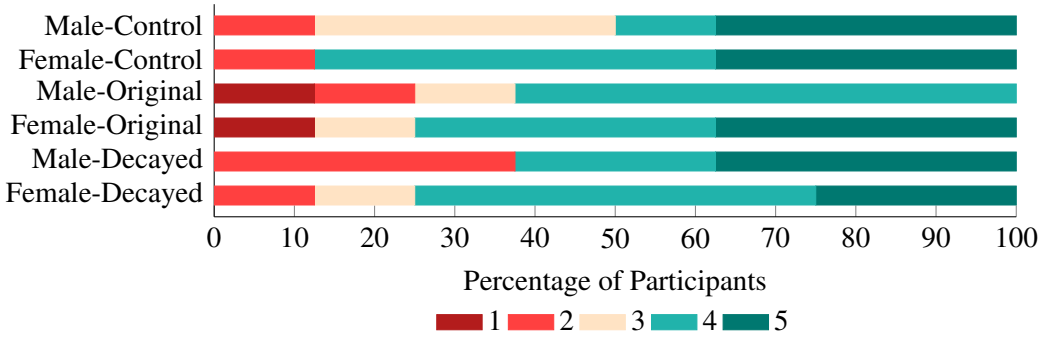


Figure 6.8: **RQ4**: Likelihood of participants looking up candidates online in real life (1 = Extremely unlikely, 5 = Extremely likely).

Descriptive statistics:

Looking up candidates online. Most of our participants look up job candidates online before hiring them, as shown in Figure 6.8. Overall, 65% of participants are likely or extremely likely to look up candidates online in real life. They also check various social media profiles for job candidates, as summarized in Table 6.7. Most participants check LinkedIn, almost half check Facebook, followed by Twitter as a close third. Some would look up Instagram or do a general Google search and check whatever comes up in their search. A couple reported looking up candidates using internal Government of Canada social media tools. Only four participants indicated that they check none of the OSNs platforms.

Table 6.8: **RQ4:** Likert-scale responses for the likelihood that participants hire a candidate who posts specific types of content in real life (n= 48, 1 = Extremely unlikely, 5 = Extremely likely). Md: median, μ: mean, and SD: standard deviation

Online content type	***Md***	μ	***SD***
Could damage company's reputation	1.0	1.4	(0.7)
Against company's values	1.0	1.4	(0.7)
Shows participation in illegal activities	1.0	1.4	(0.7)
Shows inappropriate content	2.0	1.8	(0.8)
Shows controversial content	2.0	2.5	(0.9)
Shows negative comments about past employers	2.0	2.1	(0.9)

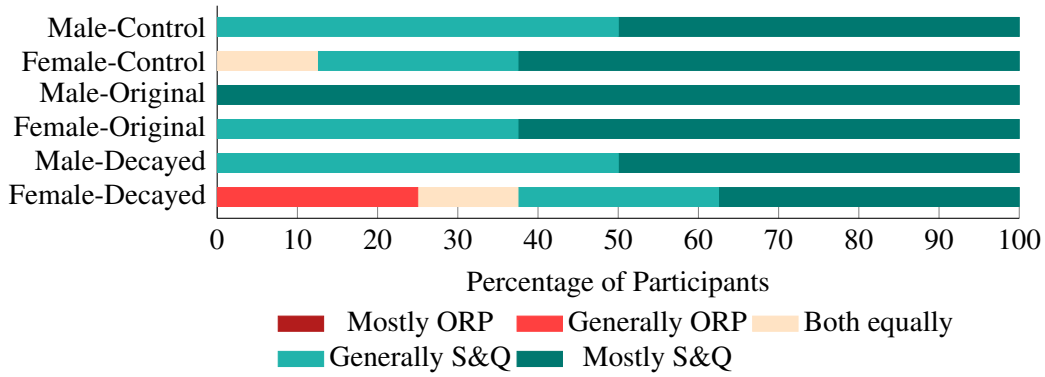


Figure 6.9: **RQ4:** Skills and qualifications (S&Q) vs. online reputation (ORP) in the hiring process in real life.

Skills & qualifications vs. online reputation. Figure 6.9 shows the most important factor identified by participants when making a hiring decision. Participants reported a candidate's skills and qualifications are more important than the candidate's online reputation. However, participants also said they were unlikely to hire a candidate who posts negative content (see Table 6.8). Responses in the figure show a similar pattern across conditions except for the female-decayed condition. We are unsure whether this difference was due to an effect of the assigned prototype (female candidate with decaying content), or whether it is due to random variance since this was a relatively small sample (8 per condition).

Timeframe. 14% of participants reported that all past online activities matter in terms of online reputation. 21% selected content from "1 year ago or less", 46% chose content posted "1-5 years ago", and 2% said that content from "5-10 years ago" matters. Finally, 17% thought that "none" of a candidate's online activities matter.

Content analysis:

We looked at responses to questions from the interview guide that considered participants' real-life practices (Section C). Those questions explored 1) whether participants would look up their candidates online, what they would look for, and how influential online reputation would be in their hiring decision, 2) if age/recency of the online content matters, 3) what kind of online activities would encourage them to hire or reject a candidate, 4) if they had ruled out or had hired a candidate because of their online reputation, and finally, 5) whether online reputation should impact someone's chances of getting a job. We extracted a total of 24 main themes from the interview scripts pertaining to these aspects and we report on each aspect individually.

1) Real-life use. Overall, most participants look up their candidate online, mainly looking for either red flags or further professional details. Most participants *always* look up their candidate online. The reasons, online resources, or OSNs platforms they consider vary based on what they are looking for. There was four main reasons for exploring social media: to look for any alarming content or *red flags*, look for further *professional details*, look for information that would *confirm their impression of the person during the interview or validate their application/CV*, and finally, look for the kind of *person they are or how they present themselves online*. Besides the group of participants who would *always* look up candidates online, a few participants either had a *limited* use, had not considered it before but *would start*, or they *never* look up their candidate online. Table 6.9 summarizes the observed frequency and reasons for the use of online reputation by participants.

Participants indicated that online reputation plays or would play a minimal role in their hiring decision unless the profile showed negative content. They thought that negative online content can disqualify the candidate even if the candidate met the professional requirements for the job. Most participants indicated that they had the additional safeguard of security checks which are done automatically as part of the hiring process in their organizations. As a result, they weigh the application and the interview more than online reputation in most cases; however, online reputation would weigh more if it triggered red flags. Overall, online reputation could sway their decision negatively against the candidate but was insufficient to sway a decision positively in favour of the candidate.

Table 6.9: **RQ4:** Main reasons participants use online reputation in the hiring process in real life. The number of participants falling under each category is indicated in each cell.

Reason	*Frequency*			
	Always	*Limited*	*Would Start*	*Never*
Red flags	13	2	3	5*
Further professional details	12	3	2	1*
Confirm or complement interview, impression, or CV	7	3	3	1*
Kind of person they are/how they present themselves online	9	2	-	1*
Identify candidates	2	-	-	-
Evidence they'd fit with the organization	1	-	-	-
How public their profiles are	1	-	-	-
Values and interests	-	-	1	-
Evidence of creativity and critical thinking	1	-	-	-
Verify the person	-	1	-	-

* Never look up candidates online but this is what they would look for if they do.

2) Relevance. Almost all participants believed that the age of content matters and that the most recent activities are the most relevant to hiring decisions. They recognized that people change or mature with time. When deciding on a candidate, this group of participants would either factor in the age of the person or factor in the level of negativity of the content itself. When factoring in the age of the person, they were typically more tolerant towards younger candidates.

On the other hand, a minority of participants thought that all past online activities are relevant because people hardly change and the past is still part of their identity.

3) Activities that encourage to hire/reject. Most participants mentioned that content showing professional engagement or work-related activities would encourage them to hire a candidate. Some participants also positively mentioned content that shows having a well-rounded lifestyle and life experience, volunteering, and community engagement.

In terms of content that could be considered extreme enough to disqualify a candidate, most participants mentioned content that is racist, extremist, misogynistic, verbally or physically violent, illegal, toxic, or hateful towards others in general. A few mentioned content with excessive/constant drinking or partying, abuse of drugs or any substance, and inappropriate or socially unacceptable behaviour. A number of participants mentioned content with negative comments against employer or company values.

4) Previous incidents. Most participants indicated that they had never come across anything that would cause them to hire or reject a candidate specifically because of their

online presence. However, a number of participants mentioned that they disregarded many applicants in the initial screening process due to their online presence. Others discussed examples of conversations they have had with their current employees about their online presence; these conversations were mainly around clarifying what is acceptable to share. A couple of participants mentioned that part of their policy is to ask candidates to delete questionable posts from their profiles before starting their new job.

5) Consequences. Most participants believed that online reputation should impact people's chances of getting a job, especially if their profiles contain negative content or if it could influence the job. Others thought it should have a limited impact or that it should affect the candidate both positively and negatively. A minority thought that it should not impact the candidate's chances but that it has nonetheless become an inevitable component to consider.

Participants generally thought the solution is for candidates to be aware of what they share, and to clean up/delete/lock down their profiles. A few participants mentioned the importance of privacy education for candidates, suggested that norms need to shift towards employers minimizing the importance of online presence, or felt that employers had a responsibility in setting expectations. For example, they thought that employers should inform candidates that their profiles will be considered, confirm findings with the candidate before deciding, or set rules and regulations around the use of online reputation in the hiring process.

RQ4 Results Summary

Both the questionnaire and interview responses showed that participants consider online reputation in their real-life practices, however, it is most influential when recent activities show extreme negative behaviour by the candidate. When checking candidates online, participants are mainly interested in looking for red flags or professional details, and are mostly focused on recent content.

6.5 Discussion

We reflect on each of our research questions and compare these results with those of the previous chapter.

6.5.1 Influence of the Decay Representation

RQ1: *How does decaying a candidate's OSN profile influence managers' hiring decisions?*

The decaying and original representations both informed participant's hiring decision and resulted in similar impressions of the candidate. Both representations led participants to note positive characteristics about the candidate and most made a positive hiring decision. However, decay representation raised fewer red flags, and participants were more likely to give the candidate the benefit of the doubt.

On the other hand, the control representation led participants to attribute fewer positive characteristics to candidates. Although hiding posts did not necessarily directly result in a negative impression, it did not help participants make a hiring decision or give an actual assessment of the candidate. It also pushed them to investigate the candidate further to find information that would help in their decision.

Results pertaining to the control representation are similar to those from the previous chapter. However, in the previous chapter, we saw a clear advantage of the decay representation over the original representation, unlike results of this study. Content of both the decaying and the original profiles did not meet the current participants' threshold for alarming content that would prevent them from hiring a candidate. They gave more weight to the most recent activities and gave candidates the benefits of the doubt. We also noted that participants seemed cautious in the interviews and generally gave more positive responses than in the anonymous survey. We discuss this further in Section 6.6.

The shrinking effect did not raise suspicion for our participants. Interestingly, participants not assigned to the decayed condition were more likely to intuitively realize the intention of the representation than those assigned to the decayed condition. This could be because it was easier to interpret in comparison to other layouts rather than as a standalone interface. Overall, most participants appreciated the idea, except for a few concerns about missing red flags from the past. These concerns somehow contradict their expressed preference for relevance of posts presented in Section 6.4.4.

6.5.2 Influence of Demographics

RQ2: *How does the gender of the candidate influence managers' hiring decisions?*
RQ3: *How do managers' demographics such as gender and age impact their hiring decisions?*

Candidates' demographics: Aligned with previous research, we saw an effect of candidate gender on responses, where a female candidate was more harshly judged when publishing online content that can be perceived as negative [66] and when it is fully visible/public [6]. As explained in the next paragraph, this effect was not based on factors such as age or gender of participants. However, it could be due to demographics not considered in our analysis, such as ethnicity/race or cultural background.

Participants' demographics: The participants' demographics we explored had limited impact on hiring decisions. Female participants rated the control (empty) profile higher when judging negative characteristics, aligning with previous research on females' perception of privacy [57].

These suggest that stereotypes against females can still exist in an employment context. The demographics effects found in this chapter are completely different than those observed in the previous chapter. We note that the number of participants in this study is significantly lower and demographics not included in our analysis (e.g., field of employment, cultural background, ethnicity) may have impacted these nuances. The fact that our participants come from Canada while participants from the previous study come from the US might have had an impact as well.

6.5.3 Use of Online Reputation

RQ4: *How do OSNs and online reputation influence managers' real-life hiring decisions?*

Online reputation influences our participants' real-life hiring practices. In total, 65% of our participants acknowledged looking up candidates online. Unlike results from the previous chapter, only about half showed interest in actively looking up personal profiles of their candidates. This could be because background security checks are performed in their departments; this is especially true in the government where these checks are extensive.

Moreover, most existing literature on the use of OSN for hiring refers to research done in a US context. Similarly, participants from our previous study were recruited from the US,

but participants in this study were recruited from Canada. It is unclear whether differences exist in the use of OSN in employment between US and Canada. Neither country has a federal or nation-wide law that is equivalent to the GDPR [54, 131], but Canada has been taking steps to conform to the GDPR [103, 104]. We also note the timing of the studies in relation to the enactment of the GDPR. Since the GDPR came into effect in May 2018, there has been an increasing attention globally around individual privacy rights and data protection laws. This may have increased participant awareness in the six months between our two studies. Furthermore, we observed that participants seemed cautious during our interviews (discussed in Section 6.6) about explicitly mentioning OSN platforms other than LinkedIn, unless we prompted them and reminded them that their responses will be pseudo-anonymized. Therefore, we suspect that the prevalence was under-reported in this study.

Personal reputation would come into play and largely inform their hiring process if profiles display red flags or content meeting their threshold of extreme negative behaviour. Online reputation does have an influence and could disqualify a professionally competent candidate.

Furthermore, participants thought that a candidate's online content is an extension of the person and it has become an inevitable component and a useful resource to inform employers' hiring decision. They believe that it is the candidate's responsibility to manage their online reputation. Unfortunately, existing issues around ORM and privacy management make it difficult for users to address their intentions and needs [5, 83, 99, 145, 147, 152].

Decay representations could plausibly address this issue; they do not raise questions or lead to assumptions from hidden content (unlike the control representation). They can automatically manage irrelevant past content (unlike the original representation). The decay representations could also help candidates more accurately and efficiently reflect the person they are today and dissociate from past content.

6.6 Limitations

The study had a usual limitation common to lab studies. We suspect the presence of social desirability bias [49] which led to participants being more cautious about their opinions face-to-face, compared to when giving responses in an anonymous online survey. Participants may have been embarrassed to reveal their own biases or hesitant to share opinions that could be against their organization's policy.

The scenario of the study was also disconnected from a real-life scenario where participants have a specific vacant position and a well-defined set of criteria for the the position to evaluate the candidate.

In addition, participants had an artificial profile presented as video while in real-life they would be able to freely navigate the profile or look up other online resources. Neither the profiles nor the study scenario explicitly stated the gender of the candidate, hence, the perceived gender of the candidate was open to participants' own definitions and interpretations. We chose not to modify any of these design choices to remain compatible with Chapter 5.

6.7 Conclusion

We extended work presented in Chapter 5 by exploring the influence of *decay representations* in a hiring context through in-person interviews. We had the same independent variables: *representation type* and *gender of candidate*. We reused the six prototypes from Chapter 5 that varied the representation of the profile and the gender of the profile owner. We conducted a 3×2 lab study with 48 participants who had a job hiring experience. We found a significant effect of *candidate's gender*, where participants who saw the original representation minimized the negative aspects of online reputation of the *male* candidate more than the *female* candidate. On the other hand, participants' demographics had a limited impact on the results. We found that both the decay and the original representations led participants to have similar impressions and better informed their hiring decision more than the control representation. Between the three representations, the decay representation helped managers attribute positive characteristics to the candidate. In addition, the decay representation did not significantly raise suspicion among participants, and they mostly

appreciated its intention. We further found that participants use online reputation in their real-life hiring practices, especially if it displays red flags. Participants believed that online activities should impact the candidate and that the responsibility is with the candidate to manage their online reputation. We believe that decay representations could plausibly help the candidate achieve such goal and dissociate from past online activities.

Chapter 7

Discussion, Conclusions, and Future Work

7.1 Thesis Summary

We return to our main research goal: exploring how OSNs can support users in managing their online reputation during important selection processes. More specifically, looking at mechanisms that could help social media users automatically dissociate from their past online activities, primarily in a hiring context. We discuss our research trajectory in addressing each of our research objectives.

Objective 1: *Understand how online reputation impacts social media users in the context of important selection processes such as a hiring process.*

In Chapter 3, we explored user perception of online reputation. Confirming previous research, we showed that online reputation could negatively influence individuals' professional lives, especially in a *political* setting compared to *employment* settings.

Objective 2: *Compare different decay representations that can help social media users automatically dissociate from their past online content.*

In Chapter 4, we compared three different representations that decay OSN content visible to others. Content shrinking and content fading best represent the aging/decaying of digital artifacts, among those studied.

Objective 3: *Identify users' attitudes and concerns towards decay representations that were presented to them.*

In Chapter 4, we showed that participants appreciated decay representations but desired complex controls to maintain the decaying process. However, users remained unlikely to perform retrospective privacy management, suggesting that automatic representations could be useful for their privacy.

Objective 4: *Evaluate how decay representations impact users' online reputation in the context of important selection processes.*

In Chapters 5 and 6, we evaluated the use of decay representations for ORM during the hiring process. Decay representations performed reasonably well compared to two other tested representations. Online reputation influences participants' real-life hiring processes, suggesting a need for tools that help users dissociate from their past online activities.

7.2 Preliminary Recommendations

Based on the literature and our findings, we provide recommendations for the privacy and HCI research community. Given that this thesis has raised additional questions and other aspects should be explored, these recommendations are preliminary in nature and intended to fuel further discussion. We also discuss further aspects related to decay representations such as feasibility, OSN platforms business model, and social acceptability.

R1: Have digital decay features enabled by default as a fail-safe mechanism and for ORM: A principle of usable security and privacy is to include the safest outcome in the path-of-least-resistance since it is likely what users will choose [120, 149]. Given that the ultimate path-of-least-resistance for users is to do nothing [149], system settings should be secure by default [120, 149]. The privacy paradox [1, 2, 5, 31, 64, 95, 101, 111, 126] also suggests that users' actions rarely match their privacy intentions. Moreover, online users are unlikely to manage their Facebook privacy settings [5, 101, 152] or their online reputation because they lack usable ORM solutions [147]. Users are also apprehensive about outright deletion of content [5, 101, 152].

Thus, fail-safe decay mechanisms could at least partially protect users from their unintended self-disclosure on public profiles. This further aligns with the Privacy-by-Design principles [22] of having preventative and default measures. We also saw in Chapters 5 and 6 that managers are actively looking at candidates' OSN profiles during the hiring process. Decay representations that gradually shrink content *can* be used for ORM. Compared to the original visibility of social media posts, decay representations led managers to better appreciate the candidates' positive traits, and positive content and focus less on their past

negative content. Therefore, the decay representations could support users in dissociating from negative past online activities.

As a result, users would be mostly relieved of the burden associated with retrospectively managing their digital artifacts. Digital decay gives temporal context to the viewer and emphasizes content that is currently most timely, indirectly supporting their online privacy by gradually removing content from the public sphere as it ages. While a user's past online activities may be positive, time-based decay shows the user's current personality, regardless of their past. Decaying would provide a path-of-least-resistance to users, which would be especially useful to those with questionable past content. It would also help users in impression management [46] with their imagined audience [46], where decay allows them to present a more recent content that potentially reflects their current personalities. We note that decay can be based on criteria other than time, as discussed in Chapter 4, but this would require more user involvement. We further discuss these criteria in Section 7.4.4.

In general, users are unlikely to perform retrospective privacy management [101, 152] and need a more seamless approach to manage past content. While decay representations do not address every issue of online reputation [147] or all aspects of online privacy, they can offer a simple tool to assist users in maintaining their privacy. Decay representations as a fail-safe mechanism could reduce possible unintended consequences resulting from long-term data availability. They can be used as a preventive mechanism requiring only minimal user engagement to support their ORM.

R2: Match the aging metaphor: Metaphors are a helpful tool that serve humans' cognitive functions [41] and metacognitive strategies [18]. Metaphors link an abstract concept to a concrete concept [41], allowing extraction of common properties from both concepts to better understand the abstract concept [41]. Metaphors have had a radical impact on interface design practices [93]. The use of metaphors in the UI can reduce the mismatch between the designer's intention and the user's mental model of the system [93].

As discussed in Section 4.5.1, participants felt that representations for aging/decay of digital artifacts should reflect the natural forgetting process. Based on our early findings, the *shrinking* and *fading* representations were found to best depict the metaphor of decaying memories [17, 118] out of three evaluated representations in this thesis. Both representations could be used either individually or potentially in combination. However, if a system

designer is faced with selecting only one approach, shrinking would be recommended since it was most preferred by participants and was thought to be most intuitive and natural. It also performed reasonably well in our more extensive testing in Chapters 5 and 6. Other research suggests that representations such as pixelation or blurring are actually ineffective at preserving privacy of social media photos [78]. In Chapter 4, we found that the pixelation representations were interpreted as "concealing"; they invoked negative connotations and aroused suspicion. Taken together, these results suggest that pixelation should be avoided as a method for increasing privacy. Online sharing and privacy are guided by complex social norms and expectations [94, 121]; any representation used should be carefully implemented to ensure that it does not inadvertently make the user appear as if they are breaking such social norms.

R3: Allow overrides: Users should be allowed to override decay defaults, if they wish. As suggested in R1, the default settings should be secure, but allowing users to have control over their content is also important. By allowing overrides, users can adjust settings and perform more selective decaying [51] and control the decay rate [10] based on the context and specific online content. This might further allow users to present a more authentic self [46, 86], or one that matches their imagined audience [86]. Whereas automating such privacy decisions may be desirable, the complex, personal, and dynamic nature of these decisions makes it unlikely that they can be performed algorithmically in a fully automated way. In particular, the risks of mis-categorization could lead to privacy violations if the user expects something to automatically decay and it does not.

Given these constraints, users should remain involved in decisions to make some content visible beyond the normal decaying period, or to avoid the decaying process altogether. It is possible that they could be assisted by the system, but the ultimate choice should rest with the user and involve a distinct, conscious decision by the user that enables reflection on their intended privacy and sharing needs. This could also support existing recommendations [11, 101, 152] suggesting that the UI should promote user reflection of aged content.

We believe our recommendations align with Principles 1, 2, 3, and 7 of the Privacy-by-Design framework [22]. Our recommendations place privacy as a core function of the user interaction (*Principle 3; privacy embedded into the design*) by reducing the long-term exposure of digital artifacts and reducing risks of privacy violations (*Principle 1; proactive*

not reactive). They seek to insert privacy into the design of OSNs by default as a fail-safe feature (*Principle 2; privacy as the default setting*). The recommendations aim to maximize privacy defaults, while giving users granular privacy options to customize their privacy preferences based on their privacy and sharing requirements (*Principle 7; keep it user-centric*). By supporting the aging metaphor, the recommendations also focus on matching users' mental models as closely as possible (*Principle 7; keep it user-centric*).

7.3 Existing Decay Mechanisms

Existing tools such as Snapchat and Facebook Stories apply decay mechanisms by making a social media post disappear after brief time. In these systems, posts are only temporarily available, usually for 24 hours. After that time limit, the post is automatically deleted or moved to the user's archive and becomes inaccessible by others.

These tools serve a related but complementary purpose in that they remove content automatically, but they are intended specifically for short lifespan posts where the user explicitly recognizes that this content should only be available briefly. On the other hand, decay representations focus on solving long-term data availability issues. For example, users might initially think that some content is harmless and post it without explicit expiry, but this content may later prove regrettable (e.g., when the content is later used by an employer to disqualify the user when applying for a job).

7.4 Implementation Considerations

We briefly consider several issues surrounding potential real-life implementation of decay representations.

7.4.1 Feasibility

Within OSNs, several implementation issues would need to be addressed when implementing decay representations. First, digital content shared on OSNs may not be exclusively controlled by its publisher/owner [136]. For example, other users may be tagged in a post, or content may be re-shared by other users. In these cases, and cases where multiparty access control is required [136], it is unclear what should happen to decaying content. Do

all instances decay at the pace set by the original owner? Should other users be able to override decay settings? What happens if content is re-posted/shared after significant time has elapsed? Does it reset to full visibility or get posted partially decayed?

Another significant concern is that traces of the digital content might still be available elsewhere outside the original OSN. For example, content may be copied or downloaded by others before the decaying process begins, leaving unaltered instances of the digital artifacts. The owner of the content may also have shared copies of content on other mediums. Thus, the feasibility of decaying social media digital artifacts might be limited when considering other aspects of online sharing and availability of online data.

We should also consider the user experience of specific user classes who may have accessibility needs. For example, by their nature, these representations would not be appropriate for user with visual impairments who use screen readers. Parallel indicators may be feasible but these are beyond the scope of this thesis.

7.4.2 OSN Platforms' Business Model

Another aspect of feasibility relates to a potential reluctance of OSN platforms to implement a decay representation/feature and incorporate it into their business models. They could fear that it would affect the user experience in general, reduce opportunities for advertising [45], or reduce opportunities for behavioural tracking[1]. However, depending on the level of granularity and how the feature is incorporated into the platform's privacy model, we believe that these could be sufficiently addressed. In addition, decay only affects copies of data accessed by viewers, not owners nor the platforms themselves, both of whom would still retain access to the original data[2].

The placement of advertising within OSNs would need consideration when incorporating decay representations: How should ads be displayed on the UI? Would they shrink along with the content? Would they maintain their size and take up more and more of the screen relative to shrinking content?

Decay representations are mainly intended for use on a user's profile page. Currently,

[1] We note that platforms have the data anyway to perform behavioural tracking.

[2] This is tangentially related to the *right to be forgotten*; but this was not the main purpose of our representations. The main focus of decay representations is on interpersonal online sharing issues.

ads are mostly placed on the homepage/timeline/news feed which would not necessarily show decayed content since they represent an aggregate of everyone's most current content. Under these circumstances, the relative size of decayed content and ads would not be relevant since they do no appear in close proximity.

7.4.3 Social Acceptability

In Chapter 4, we partially explored users' attitudes and concerns towards decay representations and the concept of digital decay in general. However, questions related to the social acceptability of decay representations remain open. For example, the following questions came up during our work. How are the representations perceived by users when applied to *their own* OSN content or that of their *peers*? Do the representations make users suspect that the owner is hiding information? Do the representations diminish the benefits of social media interaction? Do they promote privacy attitudes?

7.4.4 Decay Formula

There are several options when considering the formula specifying the rate of decay for the representations. In our studies, we explored the decay as a function of time. In its simplest form, a time-based formula would be a linear gradual decay based directly on X amount of time that has passed. It could also be time-based but exponential where the decay starts at a slow rate but then decays faster after a longer period of time. However, the decay formula can also be based on more complex criteria or user preferences. For example, it could be based on the type of content itself or the degree of relevance/importance to the user. The decay can also be selectively, only to posts fitting with users' predefined criteria while other posts remain undecayed.

7.5 Thesis Contributions

The main contributions of the research are:

1. **Chapter 3:** We examined how online users perceive several different online activities of *job* or *political* candidates and whether these activities should affect their professional lives. We launched two parallel versions an online crowdsourcing survey. One asked questions relating to evaluating job candidates/employees based on their online presence, and the second asked the same questions of political candidates. We reported results of 459 valid responses from the two parallel versions of the survey. We found that political candidates are held to higher standards than job candidates/employees. Further, the findings confirmed previous research about negative impacts on individuals' professional lives because of their digital footprint [28, 73, 87, 119, 125, 138]. The survey was the first to compare how the people's opinion differs depending on whether the candidate is vying for a job or political office.

2. **Chapter 4:** We explored the use of *decay representations* [98] on OSNs from the user's perspective. Our goal was to explore how digital decay could match users' metaphor of aging and support their privacy. We tested three different representations that decay OSN content visible to other users on three different OSN platforms through a lab study with 30 participants. We identified that the *shrinking* and *fading* representations best represent decaying of digital artifacts among those tested. We further reported participants' attitudes and concerns, and discussed their preferences regarding aging/decaying of digital artifacts. These results informed the design of the studies in Chapters 5 and 6 and also helped us frame recommendations presented in Chapter 7.

3. **Chapter 5:** We evaluated the use of *decay representations* for ORM. We examined how the application of the *shrinking* representation on a job candidate's timeline might affect managers' hiring decisions.We compared with a profile showing all posts *full size* and a third profile with *no posts*. We explored whether gender of the candidate or managers' gender or age impact their decisions. We also probed how online reputation shapes hiring decisions today. We conducted a 3×2 between-subjects online survey with 360 managers and showed them fictitious OSN profiles

for job applicants. We found that gender of the candidate had no effect and that managers' gender and age had a limited impact on the results. On the other hand, the *shrinking representation* led to significantly more positive decisions and perception of the job candidates compared to the two other representations. Our findings further showed that online reputation is influential in our managers' real-life practices, suggesting a need for tools that automatically dissociate users from past online content.

4. **Chapter 6:** We extended the study from Chapter 5 through a 3×2 between-subjects lab study. We recruited 48 participants with job hiring experience using the same prototypes and questionnaires, and an additional interview. We found that both profiles with *full size* posts and with the *shrinking representation* led to more positive assessments than the profile with *no posts*. Our profiles did not seem to meet our participants' threshold of concerning content. On the other hand, the profile with the shrinking posts led participants to attribute positive characteristics to the candidates and give them the benefit of the doubt more frequently compared to the other two representations. We found that participants were more accepting of the male candidate's negative content when content is shown in full size compared to the female candidate, but we found limited impact of participants' demographics on the results. Through qualitative analysis of the interviews, We identified the reasons why participants the use online reputation in their hiring decisions and the nuances relating to how online reputation is used. Unlike managers from Chapter 5, our participants were more conservative in their reported use of online reputation. Nevertheless, we found that it plays a role and could disqualify a candidate. Our results confirmed the need for automatic tools that dissociates users from past online activities.

5. **Chapter 7:** Extending existing literature and based on our findings, we presented recommendations for using *decay representations* as a way to more easily manage online reputation. We also discussed other aspects related to the *feasibility* of implementing the representations and of incorporating them into *OSN platforms' business model*, and *social acceptability* of the representations.

7.6 Future Work

Our work suggests several additional research directions.

Assess changes in user expectations: The survey from Chapter 3 could be repeated to see how opinions have changed given changes in the political sphere in the years since 2015[3]. Furthermore, participants may respond differently when considering specific cases. Respondents' tolerance might depend on the type of job or political position. A future study could include fictitious/real examples in the survey, and investigate responses to more concrete scenarios.

Further explore decay techniques: A future study could investigate additional decay representation techniques. The feasibility of decay representations should be assessed by exploring different scenarios for data sharing on OSN (e.g., when data has more than one owner and multi-party access control is required) and walk through possible options for implementation in each scenario. Another study could investigate possible changes in user behaviour as a result of such implementation.

An additional user study could evaluate social acceptability of decay representations. This could be a multi-session study. In a first session, a limited static snapshot of participants' social media profile could be requested. Between sessions, the profile could be manipulated by applying the decay representation. In the second session, participants could return with 1-2 acquaintances where they give their perception of their acquaintance's decaying profile.

7.7 Conclusion

This thesis focused on online reputation and mechanisms that can support users in managing their online reputation and online privacy. The literature [3, 4, 6, 8, 28, 32, 34, 48, 66, 73, 87, 119, 125, 133, 138] has shown that online reputation significantly impacts online social media users. More specifically, online reputation shapes managers' hiring decisions to the disadvantage of prospective job candidates. To diminish such negative consequences of online reputation, earlier research [76, 87, 113] in online privacy emphasized a need to design

[3] Shortly after the survey has been conducted in 2015, the liberal party won the Canadian federal elections and Donald Trump won the US presidential elections.

for *digital forgetting*. The privacy and HCI literature [11, 83, 99, 101, 145, 147, 152] also suggest that existing privacy and forgetting tools do not adequately support users' online privacy and online reputation management needs.

Aiming to address this research gap, we have investigated visual representations (i.e., *decay representations*) that can support users in dissociating from past online social media content. When applying decay representations, social media content visible to others gradually and automatically disappears over time, much like our own memories fade over time. We compared three decay representations and explored the concept of digital decay with users and early results were promising. We evaluated how decay representations impact users' online reputation in an employment scenario through a crowd-sourced survey and in-person interviews. The decay representations that apply shrinking lead to significantly more positive impressions of job candidates. Aligning with previous research, we found that users remain unlikely to perform retrospective privacy management, suggesting that these automatic representations could be useful for their privacy. Additionally, we found that online reputation is influential in hiring decisions, reinforcing a need for tools that automatically dissociate users from past online activities.

Bibliography

[1] Alessandro Acquisti and Ralph Gross. Imagined communities: Awareness, information sharing, and privacy on the facebook. In George Danezis and Philippe Golle, editors, *Privacy Enhancing Technologies*, pages 36–58. Springer Berlin, 2006.

[2] Alessandro Acquisti and Jens Grossklags. Losses, gains, and hyperbolic discounting: An experimental approach to information security attitudes and behavior. *2nd Annual Workshop on Economics and Information Security*, pages 1–27, 2003.

[3] Eric Andrew-Gee. Gaffes: When candidates accidentally tell the truth. `http://www.theglobeandmail.com/news/politics/elections/gaffes-when-a-candidate-accidentally-tells-the-truth/article26713595/`, 2015.

[4] Julio Angulo and Martin Ortlieb. "WTH..!?!" Experiences, Reactions, and Expectations Related to Online Privacy Panic Situations. In *SOUPS*, pages 19–38. USENIX Association, 2015.

[5] Oshrat Ayalon and Eran Toch. Not even past: Information aging and temporal privacy in online social networks. *Human-Computer Interaction*, 32(2):73–102, 2017.

[6] Jane Bailey, Valerie Steeves, Jacquelyn Burkell, and Priscilla Regan. Negotiating with gender stereotypes on social networking sites: From "bicycle face" to facebook. *Journal of Communication Inquiry*, 37(2):91–112, 2013.

[7] Stephen Baker. *The Numerati*. Houghton Mifflin Co., 1 edition, 2008.

[8] Susan Barnes. Privacy paradox: Social networking in the United States. *First Monday*, 11, 2006.

[9] Adam Barth, Anupam Datta, John C. Mitchell, and Helen Nissenbaum. Privacy and contextual integrity: Framework and applications. In *Proceedings of the 2006 IEEE Symposium on Security and Privacy*, SP '06, pages 184–198. IEEE Computer Society, 2006.

[10] Debjanee Barua, Judy Kay, Bob Kummerfeld, and Cecile Paris. Theoretical foundations for user-controlled forgetting in scrutable long term user models. In *Proceedings of the 23rd Australian Computer-Human Interaction Conference*, OzCHI '11, pages 40–49. ACM, 2011.

[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 *Proceedings of the 12th ACM Workshop on Privacy in the Electronic Society*, WPES '13, pages 1–12. ACM, 2013.

[12] Brenda Berkelaar and Patrice Buzzanell. Online employment screening and digital career capital. *Management Communication Quarterly*, 29(1):84–113, 2014.

[13] Andrew Besmer and Heather Richter Lipford. Tagged photos: concerns, perceptions, and protections. In *CHI Extended Abstracts*. ACM, 2009.

[14] Paul Best, Roger Manktelow, and Brian Taylor. Online communication, social media and adolescent wellbeing: A systematic narrative review. *Children and Youth Services Review*, 41:27–36, 2014.

[15] Harry N. Boone and Deborah A. Boone. Analyzing likert data. *Journal of Extension*, 50(2):1–5, 2012.

[16] Katrin Borcea-Pfitzmann, Andreas Pfitzmann, and Manuela Berg. Privacy 3.0 := data minimization + user control + contextual integrity. *Information Technology*, 53:34–40, 2011.

[17] Mark E. Bouton, James B. Nelson, and Juan M. Rosas. Stimulus generalization, context change and forgetting. *Psychological Bulletin*, 125:171–186, 1999.

[18] Leanne Bowler and Eleanor Mattern. Visual metaphors to model metacognitive strategies that support memory during the process of refinding information. In *Proceedings of the 4th Information Interaction in Context Symposium*, IIIX '12, pages 250–253. ACM, 2012.

[19] Career Builder. Overwhelming majority of companies say soft skills are just as important as hard skills, according to a new careerbuilder survey. `http://www.http://www.careerbuilder.com/share/aboutus/pressreleasesdetail.aspx?sd=4/10/2014&id=pr817&ed=12/31/2014/`, 2014.

[20] Moira Burke, Cameron Marlow, and Thomas Lento. Social network activity and social well-being. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 1909–1912. ACM, 2010.

[21] Kelly Caine, Lorraine G. Kisselburgh, and Louise Lareau. Audience visualization influences disclosures in online social networks. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, pages 1663–1668. ACM, 2011.

[22] Ann Cavoukian. Privacy by design, the 7 foundational principles: Implementation and mapping of fair information practices, 2011.

[23] Farah Chanchary and Sonia Chiasson. User perceptions of sharing, advertising, and tracking. In *Eleventh Symposium On Usable Privacy and Security (SOUPS)*, pages 53–67. USENIX Association, 2015.

[24] Jenn Chen. 15 facebook stats every marketer should know for 2019. https://sproutsocial.com/insights/facebook-stats-for-marketers/, 2019.

[25] Terence Chen, Roksana Boreli, Mohamed-Ali Kaafar, and Arik Friedman. On the effectiveness of obfuscation techniques in online social networks. In Emiliano De Cristofaro and Steven J. Murdoch, editors, *Privacy Enhancing Technologies*, pages 42–62. Springer International Publishing, 2014.

[26] Kendra Cherry. Why do we forget things? these are the 4 explanations for forgetting. https://www.verywellmind.com/explanations-for-forgetting-2795045, 2017.

[27] Jason Cheung. Ownership of content in your digital life – social media (part 1). https://digitaltattoo.ubc.ca/2018/07/09/ownership-of-content-in-your-digital-life-social-media-part-1/, 2018.

[28] Leigh A. Clark and Sherry J. Roberts. Employer's use of social networking sites: A socially irresponsible practice. *Journal of Business Ethics*, 95(4):507–525, 2010.

[29] CNN Staff. Anthony Weiner scandal: A timeline. http://www.cnn.com/2016/08/30/politics/weiner-scandal-timeline/, 2015.

[30] Chris Conley. The right to delete. In *Spring Symposium Series*, pages 53–58. AAAI, 2010.

[31] Kovila P.L. Coopamootoo and Thomas Groß. Why privacy is all but forgotten: An empirical study of privacy & sharing attitude. *Privacy Enhancing Technologies*, 2017 (4):97–118, 2017.

[32] Diane Coutu. We googled you. *Harvard Business Review*, June Issue:1–12, 2007.

[33] H. Kristel Davison, Mark N. Bing, Donald H. Kluemper, and Philip L. Roth. *Social Media as a Personnel Selection and Hiring Resource: Reservations and Recommendations*. Springer International Publishing, 2016.

[34] Vanessa A. de la Llama, Isabel Trueba, Carola Voges, Claudia Barreto, and David J. Park. At face(book) value: uses of facebook in hiring processes and the role of identity in social networks. *International Journal of Work Innovation*, 1(1):114–136, 2012.

[35] Jelle Demanet, Kristof Dhont, Lies Notebaert, Sven Pattyn, and Andre Vandierendonck. Pixelating familiar people in the media: Should masking be taken at face value? *Psychologica Belgica*, 47(4):261–276, 2007.

[36] Michael A. DeVito, Jeremy Birnholtz, Jeffery T. Hancock, Megan French, and Sunny Liu. How people form folk theories of social media feeds and what it means for how we study self-presentation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 120:1–120:12, 2018.

[37] eBizMBA Guide. Top 15 most popular social networking sites. http://www.ebizmba.com/articles/social-networking-websites/, 2017.

[38] Satu Elo and Helvi Kyngäs. The qualitative content analysis process. *Journal of Advanced Nursing*, 62(1):107–115, 2008.

[39] Chris Elsden, David S. Kirk, and Abigail C. Durrant. A quantified past: Toward design for remembering with personal informatics. *Human Computer Interaction*, 31(6):518–557, 2016.

[40] Didier Fourage and Raymond Montizaan. The impact of employer's characteristics on the willingness to hire older workers: Evidence from a stated preferences experiment. *NETSPAR Academic Series, Network for Studies on Pension, Aging and Retirement*, 2016.

[41] Dedre Gentner and Brian Bowdle. *Metaphor Processing, Psychology of.* (John Wiley & Sons, Hoboken, 2006.

[42] Daniel R George, Michael J Green, Anita M Navarro, Kelly K Stazyk, and Melissa A Clark. Medical student views on the use of facebook profile screening by residency admissions committees. *Postgraduate Medical Journal*, 90(1063):251–253, 2014.

[43] Laura Giuliano, David I. Levine, and Jonathan Leonard. Do race age and gender differences affect manager-employee relations? an analysis of quits dismissals and promotions at a large retail firm. *Working Paper Series*, 2006.

[44] Laura Giuliano, David I. Levine, and Jonathan Leonard. Manager race and the race of new hires. *Journal of Labor Economics*, 27(4):589–631, 2009.

[45] David Glance. How facebook uses the 'privacy paradox' to keep users sharing. https://theconversation.com/how-facebook-uses-the-privacy-paradox-to-keep-users-sharing-94779, 2018.

[46] Erving Goffman. *The Presentation of Self in Everyday Life*. University of Edinburgh, Social Sciences Research Centre, 1956.

[47] Scott A. Golder and Michael W. Macy. Digital footprints: Opportunities and challenges for online social research. *Annual Review of Sociology*, 40(1):129–152, 2014.

[48] S. Ryan Greysen, Terry Kind, and Katherine C. Chretien. Online professionalism and the mirror of social media. *Journal of General Internal Medicine*, 25(11):1227–1229, 2010.

[49] Pamela Grimm. *Social Desirability Bias*. American Cancer Society, 2010.

[50] Ralph Gross and Alessandro Acquisti. Information revelation and privacy in online social networks. In *Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society*, WPES '05, pages 71–80. ACM, 2005.

[51] Rebecca Gulotta, William Odom, Jodi Forlizzi, and Haakon Faste. Digital artifacts as legacy: Exploring the lifespan and value of digital data. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1813–1822. ACM, 2013.

[52] Oliver L. Haimson, Jed R. Brubaker, Lynn Dombrowski, and Gillian R. Hayes. Digital footprints and changing networks during online identity transitions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 2895–2907. ACM, 2016.

[53] Richard Harper, Eryn Whitworth, and Ruth Page. Fixity: Identity, time and durée on facebook. *Selected Papers of Internet Research (SPIR)*, IR:1–21, 2012.

[54] Kathleen Harris. How a 'right to be forgotten' could trigger a battle over free speech in canada. `https://www.cbc.ca/news/politics/privacy-freedom-expression-charter-1.4843451`, 2018.

[55] Rakibul Hasan, Eman Hassan, Yifang Li, Kelly Caine, David J. Crandall, Roberto Hoyle, and Apu Kapadia. Viewer experience of obscuring scene elements in photos to enhance privacy. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 47:1–47:13. ACM, 2018.

[56] A. Heravi, D. Mani, K-K. R. Choo, and S. Mubarak. Making decisions about self-disclosure in online social networks. In *Proceedings of Hawaii International Conference on System Sciences (HICSS)*, pages 1922–1932. ScholarSpace, 2017.

[57] Susan Herring and Sanja Kapidzic. *Teens, Gender, and Self-Presentation in Social Media*. Elsevier, 2015.

[58] Bernie Hogan. The presentation of self in the age of social media: Distinguishing performances and exhibitions online. *Bulletin of Science, Technology & Society*, 30(6):377–386, 2010.

[59] Roberto Hoyle, Robert Templeman, Steven Armes, Denise Anthony, David Crandall, and Apu Kapadia. Privacy behaviors of lifeloggers using wearable cameras. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '14, pages 571–582. ACM, 2014.

[60] Hsiu-Fang Hsieh and Sarah E. Shannon. Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9):1277–1288, 2005.

[61] Panagiotis Ilia, Iasonas Polakis, Elias Athanasopoulos, Federico Maggi, and Sotiris Ioannidis. Face/off: Preventing privacy leakage from photos in social networks. In *Proceedings of the 22Nd ACM SIGSAC Conference on Computer and Communications Security*, CCS '15, pages 781–792. ACM, 2015.

[62] Zorana Ivcevic Pringle and Nalini Ambady. Personality impressions from identity claims on facebook. *Psychology of Popular Media Culture*, 1:38–45, 01 2012.

[63] Debora Jeske and Kenneth S Shultz. Using social media content for screening in recruitment and selection: pros and cons. *Work, Employment and Society*, 30(3):535–546, 2016.

[64] Leslie K. John. *The Consumer Psychology of Online Privacy: Insights and Opportunities from Behavioral Decision Theory*. Cambridge University Press, 2015.

[65] Priit Kallas. Top 15 most popular social networking sites and apps. `https://www.dreamgrow.com/top-15-most-popular-social-networking-sites/`, 2017.

[66] Katherine Karl and Joy Peluchette. *Facebook follies: Who suffers the most?* IGI Global, 2008.

[67] Patrick Gage Kelley, Joanna Bresee, Lorrie Faith Cranor, and Robert W. Reeder. A "nutrition label" for privacy. In *Proceedings of the 5th Symposium on Usable Privacy and Security*, SOUPS '09, pages 4:1–4:12. ACM, 2009.

[68] David S. Kirk and Abigail Sellen. On human remains: Values and practice in the home archiving of cherished objects. *ACM Transactions on Computer-Human Interaction*, 17(3):10:1–10:43, 2010.

[69] Donald H. Kluemper and Peter A. Rosen. Future employment selection methods: evaluating social networking web sites. *Journal of Managerial Psychology*, 24(6):567–580, 2009.

[70] P. Korshunov, C. Araimo, F. De Simone, C. Velardo, J. L. Dugelay, and T. Ebrahimi. Subjective study of privacy filters in video surveillance. In *2012 IEEE 14th International Workshop on Multimedia Signal Processing (MMSP)*, pages 378–382, 2012.

[71] Pavel Korshunov, Andrea Melle, Jean-Luc Dugelay, and Touradj Ebrahimi. A framework for objective evaluation of privacy filters in video surveillance. *Proceedings of SPIE Volume 8856*, 8856:12, 2013.

[72] Michal Kosinski, David Stillwell, and Thore Graepel. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, 110(15):5802–5805, 2013.

[73] Nalini Kotamraju, Somaya Allouch, and Kirsten van Wingerden. *Employers' Use of Online Reputation and Social Network Sites in Job Applicant Screening and Hiring*, pages 247–269. Greyden Press, 2014.

[74] Kyung-Tag Lee, Mi-Jin Noh, and Dong-Mo Koo. Lonely people are no longer lonely on social networking sites: The mediating role of self-disclosure and social support. *Cyberpsychology, behavior and social networking*, 16:413–8, 2013.

[75] Pedro Giovanni Leon, Blase Ur, Yang Wang, Manya Sleeper, Rebecca Balebako, Richard Shay, Lujo Bauer, Mihai Christodorescu, and Lorrie Faith Cranor. What matters to users?: factors that affect users' willingness to share information with online advertisers. In *SOUPS*. ACM, 2013.

[76] Meg Leta Ambrose. Speaking of forgetting: Analysis of possible non-eu responses to the right to be forgotten and speech exception. *Telecommunications Policy*, 38:800–811, 2013.

[77] Yao Li, Alfred Kobsa, Bart P. Knijnenburg, and M-H Carolyn Nguyen. Cross-cultural privacy prediction. *Proceedings on Privacy Enhancing Technologies*, 2017(2):113–132, 2017.

[78] Yifang Li, Nishant Vishwamitra, Hongxin Hu, Bart P. Knijnenburg, and Kelly Caine. Effectiveness and users' experience of face blurring as a privacy protection for sharing photos via online social networks. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 61(1):803–807, 2017.

[79] Siân Lindley, Robert Corish, Elsa Kosmack Vaara, Pedro Ferreira, and Vygandas Simbelis. Changing perspectives of time in hci. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pages 3211–3214. ACM, 2013.

[80] Siân E. Lindley, Catherine C. Marshall, Richard Banks, Abigail Sellen, and Tim Regan. Rethinking the web as a personal archive. In *Proceedings of the 22Nd International Conference on World Wide Web*, WWW '13, pages 749–760. ACM, 2013.

[81] Yabing Liu, Krishna P. Gummadi, Balachander Krishnamurthy, and Alan Mislove. Analyzing facebook privacy settings: User expectations vs. reality. In *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, IMC '11, pages 61–70, 2011.

[82] Kate Loewenthal. An introduction to psychological tests and scales. *PsycCRITIQUES*, 42(8), 2001.

[83] Michelle Madejski, Maritza Lupe Johnson, and Steven Michael Bellovin. The failure of online social network privacy settings. Technical report, Columbia University Academic Commons, 2011.

[84] Catherine C. Marshall and Frank M. Shipman. On the institutional archiving of social media. In *Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries*, JCDL '12, pages 1–10. ACM, 2012.

[85] Catherine C. Marshall and Frank M. Shipman. An argument for archiving facebook as a heterogeneous personal store. In *Proceedings of the 14th ACM/IEEE-CS Joint Conference on Digital Libraries*, JCDL '14, pages 11–20. IEEE Press, 2014.

[86] Alice E. Marwick and danah boyd. I tweet honestly, i tweet passionately: Twitter users, context collapse, and the imagined audience. *New Media & Society*, 13(1):114–133, 2011.

[87] Viktor Mayer-Schönberger. *Delete: The Virtue of Forgetting in the Digital Age*. Princeton University Press, 2011.

[88] Alexandra Michota and Sokratis Katsikas. The evolution of privacy-by-default in social networks. In *Proceedings of the 18th Panhellenic Conference on Informatics*, PCI '14, pages 6:1–6:6, 2014.

[89] Microsoft. *Online Reputation in a Connected World*. Cross-Tab Marketing Services, Microsoft, 2010.

[90] Reham Ebada Mohamed and Sonia Chiasson. Online privacy and aging of digital artifacts. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS)*, pages 177–195. USENIX Association, 2018.

[91] Reham Ebada Mohamed, Thaís Bardini Idalino, and Sonia Chiasson. When private and professional lives meet: The impact of digital footprints on employees and political candidates. In *Proceedings of the 8th International Conference on Social Media & Society*, #SMSociety17, pages 48:1–48:5. ACM, 2017.

[92] Mainack Mondal, Johnnatan Messias, Saptarshi Ghosh, Krishna P. Gummadi, and Aniket Kate. Forgetting in social media: Understanding and controlling longitudinal exposure of socially shared data. In *Proceedings of the Twelfth USENIX Conference on Usable Privacy and Security*, SOUPS '16, pages 287–299, 2016.

[93] Dennis Neale and John Carroll. *The Role of Metaphors in User Interface Design*. Elsevier, Amsterdam, 1997.

[94] Helen Nissenbaum. *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford University Press, 2009.

[95] Patricia Norberg, Daniel R. Horne, and David Horne. The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, 41:100–126, 2007.

[96] Geoff Norman. Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, 15(5):625–632, 2010.

[97] Employment North. Top 10 qualities and skills employers are looking for. `https://employmentnorth.com/top-10-qualities-and-skills-employers-are-looking-for/`, 2015.

[98] Alexander Novotny. Signs of time: Designing social networking site profile interfaces with temporal contextual integrity. In *Human Aspects of Information Security, Privacy, and Trust*, pages 547–558. Springer International Publishing, 2015.

[99] Alexander Novotny and Sarah Spiekermann. Oblivion on the web: an inquiry of user needs and technologies. In *European Conference on Information Systems*, pages 1–15, 2014.

[100] Alexander Novotny and Sarah Spiekermann. Oblivion of online reputation: how time cues improve online recruitment. *International Journal of Electronic Business*, 13(2-3):183–204, 2017.

[101] Ayalon Ochrat and Eran Toch. Retrospective privacy: Managing longitudinal privacy in online social networks. In *Symposium on Usable Privacy and Security (SOUPS)*, pages 1–13. ACM, 2013.

[102] William Odom, Richard Banks, David Kirk, Richard Harper, Siân Lindley, and Abigail Sellen. Technology heirlooms?: Considerations for passing down and inheriting digital materials. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, pages 337–346. ACM, 2012.

[103] Office of the Privacy Commissioner of Canada. Draft OPC position on online reputation. `https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/consultation-on-online-reputation/pos_or_201801/`, 2018.

[104] Office of the Privacy Commissioner of Canada. Privacy commissioner seeks federal court determination on key issue for canadians' online reputation. `https://www.priv.gc.ca/en/opc-news/news-and-announcements/2018/an_181010/`, 2018.

[105] José Ramón Padilla-López, Alexandros Andre Chaaraoui, and Francisco Flórez-Revuelta. Visual privacy protection methods: A survey. *Expert Systems with Applications*, 42(9):4177–4195, 2015.

[106] Katharine Paljug. The personality traits that will get you hired. `https://www.businessnewsdaily.com/7950-personality-traits-hired.html`, 2018.

[107] European Parliament and Council. Directive 95/46/EC of the European parliament and of the council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. *Official Journal of the European Union L281*, 38:31–50, 1995.

[108] European Parliament and Council. 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 on the free movement of such data, and repealing directive 95/46/EC (general data protection regulation) (text with EEA relevance). *Official Journal of the European Union L119*, pages 1–88, 2016.

[109] S. Tejaswi Peesapati, Victoria Schwanda, Johnathon Schultz, Matt Lepage, So-yae Jeong, and Dan Cosley. Pensieve: Supporting everyday reminiscence. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '10, pages 2027–2036. ACM, 2010.

[110] Robert A. Peterson. A meta-analysis of cronbach's coefficient alpha. *Journal of Consumer Research*, 21(2):381–391, 1994.

[111] Chanda Phelan, Cliff Lampe, and Paul Resnick. It's creepy, but it doesn't bother me. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, CHI '16, pages 5240–5251. ACM, 2016.

[112] Vassilis Plachouras, Jochen L. Leidner, and Andrew G. Garrow. Quantifying self-reported adverse drug events on twitter: Signal and topic analysis. In *SMSociety*. ACM, 2016.

[113] Eugenia Politou, Efthimios Alepis, and Constantinos Patsakis. Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions. *Journal of Cybersecurity*, 2018.

[114] Yu Pu and Jens Grossklags. Towards a model on the factors influencing social app users' valuation of interdependent privacy. *Proceedings on Privacy Enhancing Technologies*, 2016(2):61–81, 2016.

[115] Eric D. Ragan, Hye-Chung Kum, Gurudev Ilangovan, and Han Wang. Balancing privacy and information disclosure in interactive record linkage with visual masking. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 326:1–326:12. ACM, 2018.

[116] Amon Rapp and Maurizio Tirassa. Know thyself: A theory of the self for personal informatics. *Human-Computer Interaction*, 32(5-6):335–380, 2017.

[117] Sravana Reddy and Kevin Knight. Obfuscating gender in social media writing. In *Proceedings of the First Workshop on NLP and Computational Social Science*, pages 17–26. Association for Computational Linguistics, 2016.

[118] Timothy J. Ricker, Evie Vergauwe, and Nelson Cowan. Decay theory of immediate memory: From brown (1958) to today (2014). *Quarterly Journal of Experimental Psychology*, 69(10):1969–1995, 2016.

[119] Sherry J. Roberts and Terry Roach. Social networking web sites and human resource personnel: Suggestions for job searches. *Business Communication Quarterly*, 72(1):110–114, 2009.

[120] Jerome H. Saltzer and Michael D. Schroeder. The protection of information in computer systems. *Proceedings of the IEEE*, 63(9):1278–1308, 1975.

[121] Nithya Sambasivan, Garen Checkley, Amna Batool, Nova Ahmed, David Nemer, Laura Sanely Gaytán-Lugo, Tara Matthews, Sunny Consolvo, and Elizabeth Churchill. "privacy is not for me, it's for those rich women": Performative privacy practices on mobile phones by women in south asia. In *Fourteenth Symposium on Usable Privacy and Security (SOUPS)*, pages 127–142. USENIX Association, 2018.

[122] Victoria Schwanda Sosik, Xuan Zhao, and Dan Cosley. See friendship, sort of: How conversation and digital traces might support reflection on friendships. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, CSCW '12, pages 1145–1154. ACM, 2012.

[123] Emily Shugerman. 10 students have harvard acceptances withdrawn over facebook memes. `http://www.independent.co.uk/news/world/americas/harvard-facebook-memes-student-acceptance-taken-away-withdrawn-university-a7775991.html`, 2017.

[124] Manya Sleeper, Justin Cranshaw, Patrick Gage Kelley, Blase Ur, Alessandro Acquisti, Lorrie Faith Cranor, and Norman Sadeh. "I Read My Twitter the Next Morning and Was Astonished": A Conversational Perspective on Twitter Regrets. In *CHI*. ACM, 2013.

[125] William P. Smith and Deborah L. Kidder. You've been tagged! (then again, maybe not): Employers and facebook. *Business Horizons*, 53(5):491–499, 2010.

[126] Sarah Spiekermann, Jens Grossklags, and Bettina Berendt. E-privacy in 2nd generation e-commerce: Privacy preferences versus actual behavior. In *Proceedings of the 3rd ACM Conference on Electronic Commerce*, EC '01, pages 38–47. ACM, 2001.

[127] Robert Sprague. Invasion of the social networks: Blurring the line between personal life and the employment relationship. *University of Louisville Law Review*, 50, 2011.

[128] J. William Stoughton, Lori Foster Thompson, and Adam W. Meade. Big five personality traits reflected in job applicants' social media postings. *Cyberpsychology, behavior and social networking*, 16(11):800–805, 2013.

[129] Kaveri Subrahmanyam and David Smahel. *Digital Youth: The Role of Media in Development*. Springer-Verlag New York, 2011.

[130] Lee Taber and Steve Whittaker. Personality depends on the medium: Differences in self-perception on snapchat, facebook and offline. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pages 607:1–607:13, 2018.

[131] FindLaw Legal team. Social media privacy laws. `https://consumer.findlaw.com/online-scams/social-media-privacy-laws.html`, 2019.

[132] Elizabeth Thiry, Siân Lindley, Richard Banks, and Tim Regan. Authoring personal histories: Exploring the timeline as a framework for meaning making. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1619–1628. ACM, 2013.

[133] Steven L. Thomas, Philip C. Rothschild, and Caroline Donegan. Social networking, management responsibilities, and employee rights: The evolving role of social networking in employment decisions. *Employee Responsibilities and Rights Journal*, 27(4):307–323, 2015.

[134] Daniel Trottier. Social media as surveillance: Rethinking visibility in a converging world. *Social Media as Surveillance: Rethinking Visibility in a Converging World*, pages 1–213, 2012.

[135] Henk van Rossum, Huib Gardeniers, John Borking, Ann Cavoukian, John Brans, Noel Muttupulle, and Nick Magistrale. *Privacy-Enhancing Technologies: The Path to Anonymity*. Information and Privacy Commissioner / Ontario, Canada & Registratiekamer, The Netherlands, Den Haag, 1995.

[136] Nishant Vishwamitra, Yifang Li, Kevin Wang, Hongxin Hu, Kelly Caine, and Gail-Joon Ahn. Towards pii-based multiparty access control for photo sharing in online social networks. In *Proceedings of the 22nd ACM on Symposium on Access Control Models and Technologies*, SACMAT '17 Abstracts, pages 155–166. ACM, 2017.

[137] Jessica Vitak, Pamela Wisniewski, Zahra Ashktorab, and Karla Badillo-Urquiola. Benefits and drawbacks of using social media to grieve following the loss of pet. In *Proceedings of the 8th International Conference on Social Media & Society*, #SMSociety17, pages 23:1–23:10. ACM, 2017.

[138] Margaret Vroman, Karin Stulz, Claudia Hart, and Emily Stulz. Employer liability for using social media in hiring decisions. *Journal of Social Media for Organizations*, 3:1–13, 2016.

[139] Yang Wang, Saranga Komanduri, and Pedro Giovanni. "i regretted the minute i pressed share: A qualitative study of regrets on facebook". In *Symposium on Usable Privacy and Security (SOUPS)*, pages 10:1–10:16. ACM, 2011.

[140] Yang Wang, Pedro Giovanni Leon, Alessandro Acquisti, Lorrie Faith Cranor, Alain Forget, and Norman Sadeh. A field trial of privacy nudges for facebook. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 2367–2376. ACM, 2014.

[141] Yang Wang, Pedro Giovanni Leon, Kevin Scott, Xiaoxuan Chen, Alessandro Acquisti, and Lorrie Faith Cranor. Privacy nudges for social media: An exploratory facebook study. In *Proceedings of the 22Nd International Conference on World Wide Web*, WWW '13 Companion, pages 763–770. ACM, 2013.

[142] Lyndsay A. Wasser. Social media background checks in canada – do the risks outweigh the rewards? https://mcmillan.ca/social-media-background-checks-in-Canada-do-the-risks-outweigh-the-rewards, 2013.

[143] Jason Watson, Heather Richter Lipford, and Andrew Besmer. Mapping user preference to privacy default settings. *ACM Transactions on Computer-Human Interaction*, 22(6):32:1–32:20, 2015.

[144] Lynne Webb and Nicholas Temple. *Social Media and Gender Issues*. IGI Global, 2015.

[145] Allison Woodruff. Necessary, unpleasant, and disempowering: Reputation management in the internet age. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pages 149–158, 2014.

[146] Workopolis. 14 Canadians who were fired for social media posts. https://careers.workopolis.com/advice/14-canadians-who-were-fired-for-social-media-posts/, 2015.

[147] Shuzhe Yang. Understanding the pain: Examining individuals' online reputation management behaviour and its obstacles – a grounded theory. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pages 3898–3907, 2016.

[148] Yaxing Yao, Davide Lo Re, and Yang Wang. Folk models of online behavioral advertising. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, pages 1957–1969. ACM, 2017.

[149] Ka-Ping Yee. User interaction design for secure systems. In *Proceedings of the 4th International Conference on Information and Communications Security*, ICICS, pages 278–290. Springer-Verlag, 2002.

[150] Alyson L. Young and Anabel Quan-Haase. Information revelation and internet privacy concerns on social network sites: A case study of facebook. In *Proceedings of the Fourth International Conference on Communities and Technologies*, C&T '09, pages 265–274, 2009.

[151] Xuan Zhao and Siân Lindley. Curation through use: Understanding the personal value of social media. In *SIGCHI conference on Human Factors in computing systems (CHI 2014)*, pages 2431–2440. ACM, 2014.

[152] Xuan Zhao, Niloufar Salehi, Sasha Naranjit, Sara Alwaalan, Stephen Voida, and Dan Cosley. The many faces of facebook: Experiencing social media as performance, exhibition, and personal archive. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 1–10, 2013.

Appendix A

Chapter 3 Study Material

Survey Questionnaire

1) Job Candidates

A. Demographic Questions

1. Please specify your gender:
   () F () M () Prefer not to answer
2. Please specify your age (in years): [Textfield]
3. Please specify your highest level of education:
   () No school completed () Elementary School () Post-secondary school or diploma
   () High School () Bachelor's degree () Master's degree
   () Doctorate degree () Other (specify below) [Textfield] () Prefer not to answer
4. Please specify your occupation:
   () Administrative Support (e.g., secretary, assistant)
   () Art, Writing, Journalism (e.g., author, reporter, sculptor)
   () Business, Management and Financial (e.g., manager, accountant, banker) () Education (e.g., teacher, professor)
   () Law Enforcement (e.g., police officer, probation officer) () Legal (e.g., lawyer, law clerk)
   () Medical (e.g., doctor, nurse, dentist)
   () Military (e.g., soldier, military police)
   () Science, Engineering, and IT professional (e.g., researcher, programmer, IT consultant) () Service (e.g., retail clerk, server)
   () Skilled Labor (e.g., electrician, plumber, carpenter) () Student
   () Other Professional
   () Unemployed
   () Retired
   () Other (specify below) [Textfield] () Prefer not to answer
5. Please specify your nationality: [dropdown menu of countries]
6. Please specify the average total hours you spend on the Internet daily:
   () More than 8h+ () 6-8h () 4-6h
   () 2-4h () 1-2h () less than 1h () None
7. How many social network accounts do you own? [Textfield]
8. How often have you posted online content that is:
   (1 = never, 5 = very frequently)

   - embarrassing
   - potentially damaging to your employer/company's reputation
   - against your employer/company's values and beliefs
   - racist
   - against specific groups of people
   - controversial
   - commenting on controversial subjects

- intolerant
- religious
- aggressive
- inappropriate
- includes details of your personal life
- negatively commenting about your current employer
- negatively commenting about past employers
- mentioning participation in illegal activities

B. Experience

1. Have you ever heard about people who had their past online activities analyzed during a job interview?
   () Yes () No () Prefer not to answer
   If yes, please mention how (through the newspaper, the TV, online news, personal experience, someone you know, or any other means)
2. Have you ever heard about people being fired due to their online social network content?
   () Yes () No () Prefer not to answer
   If yes, please mention how (through the newspaper, the TV, online news, personal experience, someone you know, or any other means)

C. Perception of collecting/publishing Online Social Content

1. Employers should dig deep into past online social activities as a mean of evaluating job candidates.
   Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
   ( ) ( ) ( ) ( ) ( ) ( )
2. If I have a job interview, I feel comfortable with employers analyzing my past online activities.
   Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
   ( ) ( ) ( ) ( ) ( ) ( )
3. I understand why employers analyze candidates' past online activities.
   Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
   ( ) ( ) ( ) ( ) ( ) ( )
4. Employers should fire employees who post pictures in embarrassing situations.
   Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
   ( ) ( ) ( ) ( ) ( ) ( )
5. Employees whose online activities could damage the company's reputation should be fired.
   Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
   ( ) ( ) ( ) ( ) ( ) ( )
6. Employees whose online activities are against the company's values and beliefs should be fired.
   Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
   ( ) ( ) ( ) ( ) ( ) ( )
7. Employees who express racist comments should be fired.
   Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
   ( ) ( ) ( ) ( ) ( ) ( )

8. Employees who express views against specific groups of people should be fired.

Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
( ) ( ) ( ) ( ) ( ) ( )

9. Employees who express controversial views should be fired.

Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
( ) ( ) ( ) ( ) ( ) ( )

10. Employees who comment on controversial subjects should be fired.

Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
( ) ( ) ( ) ( ) ( ) ( )

11. Employees who express intolerant views should be fired.

Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
( ) ( ) ( ) ( ) ( ) ( )

12. Employees who express religious views should be fired.

Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
( ) ( ) ( ) ( ) ( ) ( )

13. Employees who post aggressive comments should be fired.

Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
( ) ( ) ( ) ( ) ( ) ( )

14. Employees who post inappropriate content should be fired.

Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
( ) ( ) ( ) ( ) ( ) ( )

15. Employees who share details of their personal life should be fired.

Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
( ) ( ) ( ) ( ) ( ) ( )

16. Employees who comment negatively about their current employer should be fired.

Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
( ) ( ) ( ) ( ) ( ) ( )

17. Employees who comment negatively about past employers should be fired.

Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
( ) ( ) ( ) ( ) ( ) ( )

18. Employees who post about participating in illegal activities should be fired.

Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
( ) ( ) ( ) ( ) ( ) ( )

19. As an employer, I would not interview a job applicant who had inappropriate content online if it was published more than 10 years ago.

Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
( ) ( ) ( ) ( ) ( ) ( )

20. As an employer, I would not interview a job applicant who had inappropriate content online if it was published less than 5 years ago.

Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
( ) ( ) ( ) ( ) ( ) ( )

21. As an employer, I would not interview a job applicant who had inappropriate content online if it was published very recently (within the last year).

Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
( ) ( ) ( ) ( ) ( ) ( )

22. Which past online activities really matter in terms of reputation?

() None () from 1 year ago or less
() from 1-5 years ago () from 5-10 years ago () All past activities matter

D. Perception of Online Social Media Content

1. I think the online social network content of job candidates does not affect their reputation.

   Strongly agree ( ) Agree ( ) Neutral ( ) Disagree ( ) Strongly disagree ( ) Prefer not to answer ( )

2. I think people should refrain from publishing content on their online social network that might affect their future image.

   Strongly agree ( ) Agree ( ) Neutral ( ) Disagree ( ) Strongly disagree ( ) Prefer not to answer ( )

3. Should online content affect someone's ability to get a job?

   Strongly agree ( ) Agree ( ) Neutral ( ) Disagree ( ) Strongly disagree ( ) Prefer not to answer ( )

4. Does online content affect someone's ability to get a job?

   Strongly agree ( ) Agree ( ) Neutral ( ) Disagree ( ) Strongly disagree ( ) Prefer not to answer ( )

5. Should details posted by someone during non-work hours affect their professional life?

   Strongly agree ( ) Agree ( ) Neutral ( ) Disagree ( ) Strongly disagree ( ) Prefer not to answer ( )

6. Should people be able to have a public online personal life that is separate from their professional life?

   Strongly agree ( ) Agree ( ) Neutral ( ) Disagree ( ) Strongly disagree ( ) Prefer not to answer ( )

7. Should content posted before someone legally became an adult be considered by employers?

   Strongly agree ( ) Agree ( ) Neutral ( ) Disagree ( ) Strongly disagree ( ) Prefer not to answer ( )

8. Does content posted before someone legally became an adult influence employer decisions?

   Strongly agree ( ) Agree ( ) Neutral ( ) Disagree ( ) Strongly disagree ( ) Prefer not to answer ( )

E. Comments

Do you have any other comments about this subject?

2) Political Candidates

A. Demographic Questions

1. Please specify your gender:
   () F () M () Prefer not to answer
2. Please specify your age (in years): [Textfield]
3. Please specify your highest level of education:
   () No school completed () Elementary School () Post-secondary school or diploma
   () High School () Bachelor's degree () Master's degree
   () Doctorate degree () Other (specify below) [Textfield] () Prefer not to answer
4. Please specify your occupation:
   () Administrative Support (e.g., secretary, assistant)
   () Art, Writing, Journalism (e.g., author, reporter, sculptor)
   () Business, Management and Financial (e.g., manager, accountant, banker) () Education (e.g., teacher, professor)
   () Law Enforcement (e.g., police officer, probation officer) () Legal (e.g., lawyer, law clerk)
   () Medical (e.g., doctor, nurse, dentist)
   () Military (e.g., soldier, military police)
   () Science, Engineering, and IT professional (e.g., researcher, programmer, IT consultant) () Service (e.g., retail clerk, server)
   () Skilled Labor (e.g., electrician, plumber, carpenter) () Student
   () Other Professional
   () Unemployed
   () Retired
   () Other (specify below) [Textfield] () Prefer not to answer
5. Please specify your nationality: [dropdown menu of countries]
6. Please specify the average total hours you spend on the Internet daily:
   () More than 8h+ () 6-8h () 4-6h
   () 2-4h () 1-2h () less than 1h () None
7. How many social network accounts do you own? [Textfield]
8. How often have you posted online content that is: (1 = never, 5 = very frequently)

   - embarrassing
   - potentially damaging to your employer/company's reputation
   - against your employer/company's values and beliefs
   - racist
   - against specific groups of people
   - controversial
   - commenting on controversial subjects
   - intolerant
   - religious

- aggressive
- inappropriate
- includes details of your personal life
- negatively commenting about your current employer
- negatively commenting about past employers
- mentioning participation in illegal activities

B. Experience

1. Have you ever heard about political candidates who had their past online activities analyzed during elections?
   () Yes () No
   If yes, please mention how (through the newspaper, the TV, online news, personal experience, someone you know, or any other means)
2. Have you ever heard about political candidates being ruled out due to their online social networks contents?
   () Yes () No
   If yes, please mention how (through the newspaper, the TV, online news, personal experience, someone you know, or any other means)

C. Perception of collecting/publishing Online Social Content

1. Media should dig deep into past online social activities as a mean of evaluating political candidates.
   Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
   ( ) ( ) ( ) ( ) ( ) ( )
2. During elections, I feel comfortable with media analyzing candidates' past online activities.
   Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
   ( ) ( ) ( ) ( ) ( ) ( )
3. I understand why media analyzes candidates' past online activities.
   Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
   ( ) ( ) ( ) ( ) ( ) ( )
4. Political candidates who post pictures of embarrassing situations should be ruled out of elections.
   Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
   ( ) ( ) ( ) ( ) ( ) ( )
5. Political candidates whose online activities could damage the country's image should be ruled out of elections.
   Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
   ( ) ( ) ( ) ( ) ( ) ( )
6. Political candidates whose online activities are against the country's values and beliefs should be ruled out of elections.
   Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
   ( ) ( ) ( ) ( ) ( ) ( )
7. Political candidates who express racist comments should be ruled out of elections.
   Strongly agree Agree Neutral Disagree Strongly disagree Prefer not to answer
   ( ) ( ) ( ) ( ) ( ) ( )

8. Political candidates who express views against specific groups of people should be ruled out of elections.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

9. Political candidates who express controversial views should be ruled out of elections.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

10. Political candidates who comment on controversial subjects should be ruled out of elections.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

11. Political candidates who express intolerant views should be ruled out of elections.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

12. Political candidates who express religious views should be ruled out of elections.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

13. Political candidates who post aggressive comments should be ruled out of elections.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

14. Political candidates who post inappropriate content should be ruled out of elections.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

15. Political candidates who share details of their personal life should be ruled out of elections.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

16. Political candidates who comment negatively about their current political party's policies should be ruled out of elections.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

17. Political candidates who comment negatively about past political party policies should be ruled out of elections.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

18. Political candidates who post about participating in illegal activities should be ruled out of elections.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

19. As a voter, I would not elect a political candidate who had inappropriate content online if it was published more than 10 years ago.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

20. As a voter, I would not elect a political candidate who had inappropriate content online if it was published less than 5 years ago.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

21. As a voter, I would not elect a political candidate who had inappropriate content online if it was published very recently (within the last year).

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

22. As a voter, I would not elect a political candidate who had inappropriate content online if it was published before becoming a public figure.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

23. As a voter, I would not elect a political candidate who had inappropriate content online if it was published after becoming a public figure.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

24. Which past online activities that really matter in terms of reputation?

() None () from 1 year ago or less
() from 1-5 years ago () from 5-10 years ago () All past activities matter

D. Perception of the Online Social Media Content

1. I think the online social network content of political candidates does not affect their reputation.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

2. I think people should refrain from publishing content on their online social network that might affect their future image.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

3. Should online content affect someone's ability to be elected?

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

4. Does online content affect someone's ability to be elected?

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

5. Should details posted by someone during non-work hours affect their political life?

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

6. Should people be able to have a public online personal life that is separate from their political life?

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

7. Should content posted before someone legally became an adult be considered in elections?

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

8. Does content posted before someone legally became an adult influence public's voting decisions?

Strongly agree	Agree	Neutral	Disagree	Strongly disagree	Prefer not to answer
( )	( )	( )	( )	( )	( )

E. Comments

Do you have any other comments about this subject?

Appendix B

Chapter 4 Study Material

Study Tasks and Interview Questions

Part 1: Basic tasks and questions per each prototype

Study Task:

- Scroll through the displayed profile.

Example probing questions asked during or after task completion:

- Can you explain your interpretation of this visual representation of posts?
- Is such arrangement/representation of posts appealing to you?
- What do you like about such interface? / What worked well for you with this design?
- What don't you like? / What was most annoying or confusing to you?
- What would you change?
- Are any features missing?

To conclude this part of the study:

Which interface do you think is most:

- Helpful or useful
- Appealing or making sense to you?

Part 2: Interview questions (after they have used the 3 prototypes)

A. As a user browsing another friend's page:

- What was your interpretation when you saw the visual representation of posts?
- What was your reaction when you saw the visual representation?
- Did you care about seeing the original post? – when posts fade away, did that make you more curious/doubtful?
- Which technique/visual representation was more helpful in showing the decay/aging of posts?

B. As an owner of the profile:

- Would you opt for decaying/fading posts as they're getting older?
- How would you like your posts to decay, which technique was most likable to you?
- At what point, if any, would you stop caring about such artifacts/posts – when they're 1 year old? 3? 5? 10?
- In which cases do you think digital artifacts should expire/disappear? Should they expire? How? By decaying? Or by deleting forever?
- Would you prefer having the option to keep old posts the same without decaying as a way to reminiscing or highlighting a blast from the past?
- Would you want the process of decaying to be automated? Or manual? What kinds of settings would you want?
    - Select specific posts to decay based on: time of publishing, specific keywords in the caption/status, pictures taken with specific friends, posts/pictures with specific location?
- Did our study change the way you browse social media today?
- Do you think decaying can protect your online privacy? If so, which visualization from the ones you saw today would you use for privacy?

Prototype Questionnaire

Prototype A questions

Each are 5-point scales

1. The visual representation of posts easily shows that they are getting old.
( ) Strongly agree to ( ) Strongly disagree ( ) Prefer not to answer

2. The visual representation of posts was.
( ) Very meaningful to ( ) Not at all meaningful ( ) Prefer not to answer

3. The visual representation of posts was.
( ) Very confusing to ( ) Very understandable ( ) Prefer not to answer

4. The visual representation of posts was.
( ) Very complete to ( ) Missing many features that I expected
( ) Prefer not to answer

5. The visual representation of posts made me change my perspective on how I use social media today.
( ) Major change in perspective to ( ) No change in perspective
( ) Prefer not to answer

6. The aging technique used in the posts was.
( ) very appropriate for the content to ( ) did not apply to the content at all
( ) Prefer not to answer

7. The visual representation of posts was.
( ) Very obtrusive to ( ) Not at all obtrusive ( ) Prefer not to answer

8. The visual representation of photo posts was intuitive to me.

( ) Very intuitive to ( ) Not at all intuitive ( ) Prefer not to answer

9. The visual representation of text posts was intuitive to me.

( ) Very intuitive to ( ) Not at all intuitive ( ) Prefer not to answer

10. If available, I would choose to use this visual representation for my social media account.
( ) Strongly agree to ( ) Strongly disagree ( ) Prefer not to answer

Prototype B questions
[Same questions above to be copied]

Prototype C questions
[Same questions above to be copied]

Wrap-up Questionnaire

How necessary is aging of posts in social media?
( ) Very necessary to ( ) Very unnecessary ( ) Prefer not to answer

If available, would you choose to have your posts age? Why or why not?

Can you describe a situation where aging of posts would have been particularly beneficial to you?

Can you describe a situation where aging of posts would have been particularly problematic for you?

Appendix C

Chapter 5 Study Material

A. Demographic Questions:

Please specify your gender.

() F () M () Other(specify below) [Textfield]

() Prefer not to answer

Please specify your age (in years). [Textfield accepting only numbers]

Please specify your highest level of education completed or in-progress.

() Post-secondary school or diploma () High School

() Bachelor's degree () Master's degree

() Doctorate degree

() Other (specify below) [Textfield]

() Prefer not to answer

Please specify in which field you are a manager.

() Agriculture () Art () Business () Communication

() Education () Engineering () Environment () Financial

() Government () Health Care () Legal () Manufacturing & Production

() Service Sector () Technical () Technology: computers

() Technology: non-computers () Other (specify below) [Textfield]

() Prefer not to answer

Please specify your country of residence.

() US () Other (specify below) [Textfield] () Prefer not to answer

Please specify your job title. [Textfield]

Please specify how many years of experience you have in your managerial position. [Textfield]

Please list example duties you perform in your position. [Textfield]

Please indicate the average number of hiring decisions you make per year. [Textfield]

How familiar are you with online social networks (e.g., Facebook, LinkedIn)?

() Very familiar () Moderately familiar

() Somewhat familiar () Slightly familiar

() Not at all familiar () Prefer not to answer

B. Ranking person's qualities/traits:

On a scale of 1 to 5, where 1 is not important at all and 5 is very important:

How important it is that a job candidate is:

- Self-motivated
- Hard working
- Loyal
- Dependable
- Team-oriented
- Confident
- Adept communicator
- Respectful to work ethics
- Flexible
- Cautious

Ranking person's online reputation:

On a scale of 1 to 5, where 1 is not important at all and 5 is very important:

How important it is that the candidate's online reputation shows:

- Positive lifestyle
- Positive attitude
- Professionalism
- Volunteering and charitable giving
- Openness and willingness to undertake new endeavours

How important it is that the candidate's online reputation avoids:

- A history of using drugs or drinking alcohol
- Offence or criticism to past employer/coworker
- Inappropriate photos or videos
- Inappropriate comments
- Controversial content

Play and watch the video

C. Video verification questions:

What is the name of the owner of the profile shown in the video?

() Fred Rayerson () Connor Bart () Dave McNeil [Male condition]

() Fanny Rayerson () Lilly Bart () Diane McNeil [Female condition]

What was the most recent status update on the profile?

() Travelling is one of the best pleasures in the world.

() One of the best holidays ever!

() Any recommendations for a recent top-selling novel?

Where did this person go on vacation?

() Florence () Milan () Verona

Where was this person employed before?

() Nokia () Blackberry () Ericsson

Was the oldest post on the profile: a textual status update or an image upload?

() textual status () image upload

D. Hiring decision based on the video:

Assume that Connor Bart has applied for a job. He possesses all knowledge, skills, and qualifications required for the job. You looked him up online and you came across their Facebook profile shown in the video and you are deciding whether to hire him.

Text in [ ] in the following questions indicates rewording in the Female condition:

How likely are you to hire Connor [Lilly] Bart?

Very likely	Somehow likely	Neutral	Somehow unlikely	Very unlikely
( )	( )	( )	( )	( )

What most influenced your decision to hire or not hire Connor [Lilly] Bart?

[Textfield]

How likely are you to do more research about Connor [Lilly] Bart's previous online activities before making a final hiring decision?

Very likely	Somehow likely	Neutral	Somehow unlikely	Very unlikely
( )	( )	( )	( )	( )

E. Evaluation of the person against the 20 criteria:

Rate the following on a scale of 1 to 5, where 1 is Strongly Disagree and 5 Strongly Agree. The content of the profile in the video shows that Connor [Lilly] Bart is:

- self-motivated.
- hard working.
- confident.
- an adept communicator.
- respectful to work ethics.
- flexible.
- cautious.

Rate the following on a scale of 1 to 5, where 1 is Strongly Disagree and 5 Strongly Agree. Connor [Lilly] Bart's online reputation shows:

- a history of using drugs or drinking alcohol.
- offence or criticism of past employer/coworker.
- inappropriate photos or videos.
- inappropriate comments.
- controversial content.
- a positive lifestyle.
- a positive attitude.
- professionalism.
- volunteering or charitable giving activities.
- openness and willingness to undertake new endeavours.

F. Visual Representation and Privacy:

It was easy to see that the posts were getting older as the video scrolled down the timeline.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree
( )	( )	( )	( )	( )

Differentiating between older and more recent posts was:

Very easy	Easy	Neutral	Difficult	Very difficult
( )	( )	( )	( )	( )

The owner of this profile values their online privacy.

Strongly agree	Agree	Neutral	Disagree	Strongly disagree
( )	( )	( )	( )	( )

Did the visual representation of the posts on the timeline influence your hiring decision?

Strong positive Influence ( ) Somehow positive ( ) Neither positive nor negative ( )

No influence at all ( ) Somehow negative ( ) Strong negative Influence ( )

G. Hiring decisions in real-life practices (Irrelevant to the video):

How likely are you to look up job candidates online before hiring them?

Very likely ( ) Somehow likely ( ) Neutral ( ) Somehow unlikely ( ) Very unlikely ( )

How likely are you to look up social media profiles of job candidates before hiring them?

Very likely ( ) Somehow likely ( ) Neutral ( ) Somehow unlikely ( ) Very unlikely ( )

How does online reputation influence your hiring decisions in your real-life practices?

[Textfield]

What is the most important factor in your hiring decision?

Mostly online reputation ( ) Generally online reputation ( ) Both equally ( )

Generally skills and qualifications ( ) Mostly skills qualifications ( )

A person's entire online history is relevant for hiring decisions.

Strongly agree ( ) Agree ( ) Neutral ( ) Disagree ( ) Strongly disagree ( )

Content posted before someone legally became an adult is relevant for hiring decisions.

Strongly agree ( ) Agree ( ) Neutral ( ) Disagree ( ) Strongly disagree ( )

Which past online activities really matter in terms of reputation when making hiring decisions?

() None () From 1 year ago or less () From 1-5 years ago

() From 5-10 years ago () All past activities matter

On a scale of 5 to 1, where 5= very likely and 1= very unlikely:

How likely are you to hire a candidate whose online activities:

- could damage the company's reputation
- are against the company's values and beliefs
- show participation in illegal activities
- show inappropriate content
- show controversial content
- show negative comments about past employers

Do you have any more thoughts or comments that you would like to share with us?

[Textfield]

Appendix D

Chapter 6 Study Material

Interview questions

A- Experience in general:

1. For how long have you been making hiring decisions?

2. How often do you have to make a hiring decision?

3. What qualifications and personal qualities do you look for in a candidate?

4. How much information do you expect to see on someone's social media profile?

    a. How about their Facebook or Twitter? [If they don't mention them]

Introducing the study scenario and video:

I want you to imagine that you have a job vacancy. You have already advertised the job description and interviewed potential candidates and found a qualifying candidate for the job. The candidate possesses all knowledge, skills, and qualifications required for the job. Now, you want to look up your candidate online and you come across their Facebook profile.

I will ask you to watch this video, then I will follow up with few questions. The video scrolls through the candidate's Facebook profile.

It has no audio, and you can pause, move forward or backward, up to you. Just let me know when you're done watching.

B- Follow-up questions:

5. Would you hire the candidate?
    a. Why? /Why not?

6. What influenced your decision?

7. With respect to personal qualities we talked about earlier, how does this person meet your criteria??

8. What did you think of the candidate's online activities?

9. Based on what you have seen in the video, how concerned is the person about their privacy?

10. How could you differentiate between older and most recent posts?
    [Decayed, Undecayed]

11. What would it mean to you if you found no posts or very few posts shown? [Decayed, Undecayed]

    What does it mean to you to find no posts or if you found very few posts shown? [Control]

[Decayed condition only:]

12. Pointing out the shrinking effect: If you saw this representation, what does this mean to you? / what does it tell you?

13. How would this visual representation influence your hiring decision?

14. The intention of this is to put less attention on the older posts by making them smaller in size, what are your comments about this?

C - Hiring Experience relating to online reputation:

15. Do you usually look up your potential candidates online?
    a. Why? or why not?
    b. What do you look for?

16. What about other sources like Facebook, Instagram, Twitter? What about a Google search? [If they only mention LinkedIn]
    a. What do you look for?

17. How much do you trust online content as representative of the person?

18. How does online reputation/activities influence your hiring decision?

19. How does the age of the online activity matter?
    a. Do past online activities become irrelevant after a time?
    b. How do you decide?

20. How do you weigh the candidate's qualifications against their online reputation?

21. What type of online activities could encourage you to hire a candidate?

22. What type of online activities could encourage you to reject the candidate?

23. Have you previously denied employment to a qualified candidate the job because of their online activities?
    a. Why or why not?
    b. How often does that happen?

24. Have you previously hired a candidate for a job despite negative online activities?
    a. Why or why not?
    b. How often does that happen?

25. Do you think online activities should impact someone's chances of getting a job?
    a. What do you think might be a solution to this issue?

26. Do you have any more comments you would like to share with us?

[Undecayed/Control conditions: Showing participants the decayed video after filling in the online questionnaire and following up with the following questions:]

27. Pointing out the shrinking effect: If you saw this representation, what does this mean to you? / what does it tell you?

28. How would this visual representation influence your hiring decision?

29. How would you differentiate between older and most recent posts?

30. The intention of this is to put less attention on the older posts by making them smaller in size, what are your comments about this?