When private and professional lives meet: The impact of digital footprints on employees and political candidates

Reham Ebada Mohamed
Carleton University
Ottawa, Ontario
riham.mohamed@carleton.ca

Thais Bardini Idalino
University of Ottawa
Ottawa, Ontario
tbard069@uottawa.ca

Sonia Chiasson
Carleton University
Ottawa, Ontario
chiasson@scs.carleton.ca

ABSTRACT
We present the results of a between-subjects survey with 459 participants to gather opinions of privacy and how such online content should impact job candidates and political candidates, respectively. Our analysis explores differences between the two scenarios, and whether demographic characteristics influence users' perspectives towards politicians and/or employees. Overall, respondents were less tolerant of the online activities of political candidates. We conclude the paper with a discussion of how the concept of online privacy is evolving in this age of social media.

CCS Concepts
• Information systems □ Database management system engines • Computing methodologies □ Massively parallel and high-performance simulations.

Keywords
Survey; digital footprint; job and political candidate.

1. INTRODUCTION
Most online users regularly share information on social media. Once published, it is hard to control who can access, copy, or modify such information. Additionally, the private versus public boundaries of social media spaces are unclear, thus posting online can have unintended consequences [4]. Some consequences can present benefits to society, such when criminals are found due to their online activities. However, online data can also be used to group/label people, which can bring undesirable consequences [3]. Moreover, it may affect those seeking employment or political candidates seeking election. This survey explores users' perception of various online activities and whether online reputation should be considered by potential employers and the voting public. We examine which types of posted content people find unacceptable, whether recency of the content is relevant, and whether someone's past online behaviour should impact their professional/political life.

We launched two parallel versions of the survey. One asked questions relating to evaluating job candidates based on their online presence, and the second asked the same questions of political candidates. We collected valid responses from 459 participants. Our results suggest that political candidates are held to higher standards than job candidates. To our knowledge, this work is the first to compare how the people's opinion differs depending on whether the candidate is vying for a job or political office. This work also partially supports the findings from previous work [11] showing that many people have had their career compromised by their online activities and online footprint.

2. RELATED WORK
"Digital shadows" or "digital footprints" relate to the traces of information that we produce every day and the concern that arises on who can access and what can be done to this information [11]. There are some societal benefits arising from digital footprints. For example, we can use this information to study human behaviour and social interactions [8, 13].

However, we are fundamentally changing how people interact with each other by keeping these digital footprints. Many information systems neglect the fact that humans' ability to forget is a mechanism to put past events in temporal perspective [11]. Aylon and Toch [12] found a negative correlation between the time in which information was published on Facebook and its owner's sharing preferences, indicating that users would prefer that others forget older content.

Sharing information on social media is so common that people rarely think about the consequences before posting. Additionally, social media users want to express themselves without being anonymous and use sites like Facebook as a means of broadcast [6]. Studies show that many users later regret sharing content online. As noted by Wang et al. [15], 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 [14]. These regrets were related to the repercussion of the posts in their own lives, which sometimes caused irreversible damage [2, 9]. Several studies have also explored the unintended consequences resulting from personal data shared by others, such as through being tagged in social media photos [5].

Repercussions are not limited to personal consequences. For example, the reputations of health professionals have been impacted as patients search for the digital footprint of their practitioners to evaluate their professionalism [9]. Mayer-Schönberger [11] has also documented several other cases where people have had their professional career compromised by online content.

SMSociety'17, Toronto, ON, Canada

DOI: 10.1145/3097286.3097334

Although research has been conducted on related topics, 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.

3. METHODOLOGY

We were most interested in the following two questions:

R1: Do participants hold political candidates and job candidates to different standards with respect to online content?

R2: Does participants' age/gender/nationality impact views?

The study was cleared by our Research Ethics Board.

3.1 Structure of the 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. We collected the participants' gender, age, nationality, education level, occupation, hours spent online, and number of online accounts.

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

D. 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/vote against 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).

E. 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).

Each question also 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 the verification questions before analysis.

3.2 Participants

Data was collected between November 2015 and April 2016. We launched the surveys through Crowdflower, 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 to workers from Canada, US, and UK only since the political process in these countries is somewhat similar. 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) but after validation of the Crowdflower responses and removal of surveys with missing data, we had 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. As a result, we had uneven numbers per condition. Table 1 lists participant demographics. Most participants had at least a high-school diploma and were moderate to active online users.

Table 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

4. ANALYSIS

In this section, we summarize our participants' demographics and 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's behaviour) and 5 to the most positive response (i.e., were more accepting of the candidate's behavior).

We used Mann-Whitney U test to compare factors (described below) between the two surveys. To examine if age, gender, or nationality influenced responses, we used Ordinal Regression, Mann-Whitney, and Kruskal-Wallis respectively. Tests assume a significance level of $p < 0.05$ unless otherwise noted.

4.1 Factor Analysis (FA)

As in earlier work [10], we conducted Principal Component Factor Analysis to identify the most significant variables (questions) and group them into factors based on similar response patterns. We conducted this analysis on each survey separately, considering the 29 questions from Sections C and D of the survey to reduce the questions to a smaller number of variables for subsequent analysis. As expected, the resulting Rotating

Component Matrix (showing how variables are grouped into factors) was different in each survey. We ignored correlations of less than 0.5. We extracted groupings common across both surveys, resulting in 19 questions grouped into seven factors. The descriptions are phrased in terms of the job survey, with modifications for the political survey in parentheses.

F1: Content Investigation. Questions C1 and C2 asked if respondents were comfortable with having past online activities of candidates investigated by employers (media).

F2: 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 company's (country's) reputation, (ii) went against the company's (country's) values, (iii) were illegal.

F3: 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).

F4: 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.

F5: 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.

F6: Time Span. Questions C19, C20, and C21 explored whether 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.

F7: Consequences. Questions D3 and D5 explored whether online content (even when posted during non-work hours) should affect candidates' ability to get a job (be elected).

Table 2. Mean values per factor in each survey.

Factors	Job	Political
F1: Content investigation	3.0	**2.7**
F2: Reputation	2.7	2.6
F3: Offending employers/parties	**3.3**	3.7
F4: Controversial content	4.0	**3.7**
F5: Discriminatory content	2.8	**2.6**
F6: Time span	3.4	**2.9**
F7: Consequences	3.2	**2.8**

The factor analysis was generally as expected, the questions included in each factor were thematically related and could plausibly be grouped. One unexpected grouping was F4, where questions relating to personal details and controversial data were grouped under one factor. We believe this may be because some participants interpreted the term 'personal details' as 'intimate' or otherwise more 'controversial' than we had initially intended.

Reliability Analysis tests on the factors were used to confirm how closely each set of questions are related as a group. Cronbach's alpha for all factors in both surveys was above 0.7, with the exception of two factors in the job survey (0.62 and 0.65). Following the reliability test, we computed a factor score by calculating the average value of included questions. These factor scores were used for the further analysis.

4.2 Job vs. Political Candidates (R1)

We address our first research question, R1, by comparing the responses to the job and the political surveys using Mann-Whitney tests on the seven factors, with Bonferroni correction. Table 2 summarizes the mean values for each factor. In this, and subsequent tables, lower means are in bold-red and pairs where significant differences were found have grey backgrounds. We found a significant difference in the following six factors. For five factors, respondents were less tolerant of political candidates than job candidates.

F1: Content Investigation. Political survey respondents were more inclined to investigate the online content of candidates than those assessing job candidates ($p=0.001$).

F3: 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 ($p=0.000$).

F4: 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 ($p=0.000$).

F5: 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 ($p=0.002$).

F6: 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 whose inappropriate online content had been published up to 10 years ago than those evaluating job candidates ($p=0.000$).

F7: Consequences. Political survey respondents believed more strongly that online content of candidates should affect their professional life and ability of being elected compared to those assessing job candidates ($p=0.000$).

4.3 Effect of Age, Gender, Nationality (R2)

To address research question R2, we investigated the effects of age, gender, and nationality on responses.

Age. We used Ordinal Regression to explore whether age affected the seven factors. Results showed no significance on any factor within the job survey. However, there was a significant effect of age on three factors in the political survey; *F2: Reputation* ($p=0.001$), *F3: Offending employers/parties* ($p=0.003$), *F4: Controversial Content* ($p=0.000$). Using odds ratios (OR) based on the beta values (estimates), we found that older participants were more tolerant of political candidates posting content that might *threaten the image* of the country (OR=0.96). However, they were less tolerant of political candidates posting content that might *offend* (OR=1) their political party, or be considered *controversial* (OR=1).

Gender. We used independent samples Mann-Whitney tests to examine if gender affected the seven factors. Mean values for each factor per gender are shown in Table 3. In the job survey, we found no significant effect of gender, except for *F2: Reputation* ($p=0.026$). 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 *F2: Reputation* (p=0.011) and *F4: 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 factors. Mean values per country are shown in Table 3. A significant effect of nationality was found on both surveys for two factors: *F4: Controversial Content* (Job: p=0.003, Political: p=0.000), and *F5: Discriminatory Content* (Job: p=0.023 and Political: p=0.007). Additionally, a significant effect was found on *F3: Offending Employers/Parties* on the political survey (p=0.000). Canadians were less tolerant of employees and political candidates who post controversial content. Moreover, they were less tolerant of political candidates who post content offensive to their political party. And finally, responses from the UK were less tolerant of employees who post discriminatory content, while responses from the US were less tolerant of political candidates who post such discriminatory content. This is especially interesting given the 2017 political climate in the US.

Table 3. Mean values per factor for each gender and nationality, with significant results highlighted.

	Gender/ Nationality	F1	F2	F3	F4	F5	F6	F7
Job	**F**	3.0	**2.6**	3.2	4.0	2.8	3.4	3.2
	M	3.0	2.8	3.4	3.9	2.8	3.4	3.3
Political	**F**	2.7	**2.4**	3.8	3.9	2.4	2.9	2.7
	M	2.6	2.7	3.7	**3.6**	2.6	2.9	2.8
Job	**Canada**	3.1	2.8	3.3	**3.8**	2.9	3.4	3.4
	UK	3.0	2.6	3.3	4.1	**2.6**	3.6	3.3
	US	3.0	2.6	3.3	4.0	2.9	3.4	3.1
Political	**Canada**	2.8	2.6	**3.4**	**3.5**	2.6	2.9	2.8
	UK	2.6	2.8	4.2	4.2	3.0	3.1	2.9
	US	2.6	2.4	3.8	3.8	**2.3**	2.8	2.7

4.4 Participants' Own Online Behaviour

To put participants' responses into context, we also asked them about their own habits in posting online content. Figures 1 and 2 summarize their responses. As shown, participants in both groups reported similar posting habits.

5. 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 job candidates (R1). For R2, we found that older respondents were less tolerant of political candidates' online activities on two of seven factors and more tolerant on a third factor. Gender also impacted respondents' opinion of political candidates' online activities, but no clear pattern emerged. Nationality impacted responses on three of seven factors, with North American respondents being less tolerant. There were far fewer differences for on the job candidates survey.

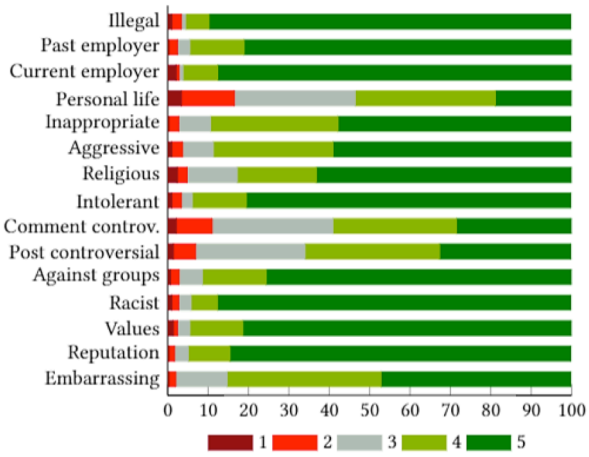

Figure 1. Job Survey: Frequency of posting specific categories of online content (1 = very frequently, 5 = never).

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

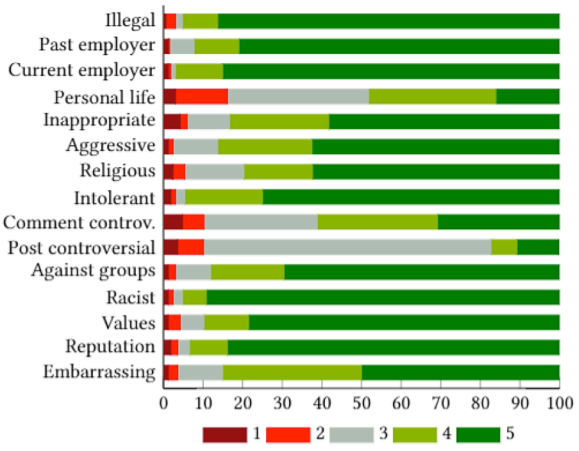

Figure 2. Political Survey: Frequency of posting specific categories of online content (1 = very frequently, 5 = never).

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 role models, 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 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 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 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 believe that this survey raises interesting questions about social medial, privacy, and digital footprints. As more of our lives are digitized, we, as a society, need to consider the implications. It is un- likely 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 were social media content may have short-term 'good' consequences, it can lead to 'bad' long-term outcomes for users. On the other hand, it makes it easier for others to assess a candidate's character, assuming that posted content is legitimate.

6. LIMITATIONS AND FUTURE WORK

Although generally accepted within the usable security and HCI communities, crowd-sourcing 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.

We will complete qualitative analysis of the open-ended questions from the surveys to gain further insight into respondents' views. Additionally, we will further examine whether participants' online behaviour had an impact on their perception of job/political candidates' social media content. We will also explore repeating the survey to see how opinions have changed over the last few months. Furthermore, participants may respond differently when considering specific cases. For example, they may ignore their values when voting for a specific candidate. Moreover, respondents' tolerance might depend on the type of job or the political position of the candidate. A future study could include fictitious/real examples in the survey, and investigate responses to more concrete scenarios.

7. CONCLUSION

We conducted a survey investigating how social media activities and digital footprints affect job and political candidates. We found significant differences in how these two types of candidates were evaluated by respondents. Respondents' demographic factors had limited impact on the results. Results reflect incidents reported in the media where some politicians had their career compromised by their digital footprint [1, 7], and research on the impact of online content on professional careers [11]. 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.

8. ACKNOWLEDGMENTS

T. Bardini Idalino acknowledges funding granted from CNPq-Brazil [233697/2014-4]. R. Mohamed acknowledges funding granted from the Ontario Trillium Scholarship (OTS) program. S. Chiasson acknowledges funding from NSERC for her Canada Research Chair and Discovery Grant.

9. REFERENCES

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

[2] Julio Angulo and Martin Ortlieb. 2015. "WTH..?!" Experiences, reactions, and expectations related to online privacy panic situations. In *SOUPS*. USENIX Association.

[3] Stephen Baker. 2008. *The Numerati* . Houghton Mifflin Co.

[4] Susan Barnes. 2006. Privacy paradox: Social networking in the United States.

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

[6] Angela M. Cirucci. 2015. Facebook's affordances, visible culture, and anti-anonymity. In *SMSociety*. ACM.

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

[8] Scott A. Golder and Michael W. Macy. 2014. Digital footprints: Opportunities and challenges for online social research.

[9] S. Ryan Greysen, Terry Kind, and Katherine C. Chretien. 2010. Online professionalism and the mirror of social media.

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

[11] Viktor Mayer-Schönberger. 2011. *Delete: The virtue of forgetting in the digital age.* Princeton University Press.

[12] Ayalon Ochrat and Eran Toch. 2013. Retrospective privacy: managing longitudinal privacy in online social networks. In *SOUPS*. ACM.

[13] Vassilis Plachouras, Jochen L. Leidner, and Andrew G. Garrow. 2016. Quantifying self-reported adverse drug events on Twitter: Signal and topic analysis. In *SMSociety*. ACM.

[14] Manya Sleeper, Justin Cranshaw, Patrick Gage Kelley, Blase Ur, Alessandro Acquisti, Lorrie Faith Cranor, and Norman Sadeh. 2013. "I read my Twitter the next morning and was astonished": A conversational perspective on Twitter regrets. In *CHI*. ACM.

[15] Yang Wang, Saranga Komanduri, and Pedro Giovanni. 2011. "I regretted the minute I pressed share: A qualitative study of regrets on Facebook". In *SOUPS*. ACM.