Enhancing SNS Profile Writing with a Search-Based Assistant System

Kousuke Nagase[1], Hideo Joho[1]

[1]University of Tsukuba, Tennodai 1 1 1, Tsukuba, Ibaraki, Japan

Abstract

In Social Networking Services (SNS), user profiles, which often consist of an image, texts, and other items, have an important role to connect with other users. However, in a preliminary study with 3,193 sample profiles on Twitter, we found that the average length of profile texts was 40 characters ($SD = 32.52$) where the maximum length is 160. This suggests that many SNS users are missing potential opportunities to expand their social network due to the short profile texts. Therefore, we proposed a search-based interactive system to support the writing of profile texts in SNS. The proposed system was designed to dynamically search for similar profile texts while users were typing their profile so that they can get new ideas such as what to write or how to express. We evaluated the effect of the proposed system by a user study with 24 participants, and found that the proposed system enabled participants to significantly increase the length of written profile texts (+92.5% on average), compared to a baseline system with no assistance.

Keywords

Social Network, User Profile, Writing Assistant, Dynamic Search, User Study

1. Introduction

1.1. Research Background

Social Networking Services (SNS) have exploded in popularity. According to a survey by Metaxas et al. (2014), self-expression and networking were the two main purposes for using Twitter with the proportion of 35.7% and 33.2%, respectively [1]. For the self-expression and networking on SNS, the user profile is essential to establish new connections with other users [2, 3]. For example, a study reported that longer self-description texts were perceived to be more trustworthy [4]. In an experimental study by Counts et al. (2009), “Quotes” and “About” were found to be useful in expressing personality traits [5].

However, many SNS users do not have a rich profile text. Figure 1 shows the distribution of the number of characters in the profile text of 3,193 Twitter accounts who are deemed to associate with one of the major universities in Japan. The accounts were manually collected by one of the authors from Twitter lists that were distributed by the associated members of the University. Figure 1 shows that a large proportion of accounts uses less than half of the maximum number of characters which is 160. The average length of profile texts was 40. Even though some of them wrote limited information intentionally due to some reasons like a privacy concern [6, 7], there might be many users who can benefit from enriching their profile text in social networking. This observation led us to develop and evaluate a mechanism that enables SNS users to generate richer profile texts.

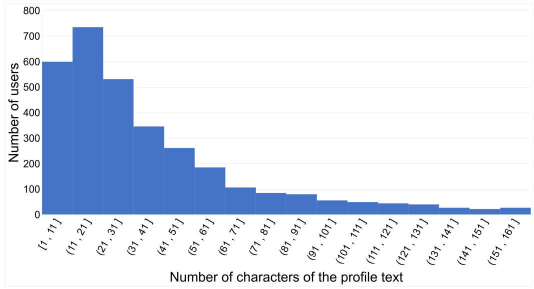

Figure 1: Distribution of the number of characters in the profile text of Twitter accounts ($N = 3,193$, $Mean = 40$, $SD = 32.52$)

The current work was inspired by the concept of Observational Learning [8] in the field of Social Psychology. In a broad sense, Observational Learning is the process of learning something by observing the behaviour of others. We thought that the behaviour of referring to the profile text of other users could be regarded as a form of observational learning.

1.2. Related Works

Research related to the present study includes writing support for different types of texts, which consist of novels, scientific paper, sentences written in a foreign language [9, 10, 11]. Roemmele, et al. developed a system to support the creation of story texts by suggesting the

DESIRES 2021 – 2nd International Conference on Design of Experimental Search Information REtrieval Systems, September 15–18, 2021, Padua, Italy
nagase.kosuke.sw@alumni.tsukuba.ac.jp (K. Nagase); hideo@slis.tsukuba.ac.jp (H. Joho)

completion of the next sentence from a previously created text [9]. Once the user has decided on the theme and vision of the story, he or she can write specific sentences with the help of the program. Kinnunen et al. developed a system that checks for criteria such as "are keywords that occur frequently in the abstract also used in the title" to help users write readable and consistent scientific papers [10].

Ellison, et al. [2] conducted an interview study with users of an online dating service and found that many participants used other users' profile information to find out what they should pay attention to in constructing their profiles. For example, one participant stated that she avoided using a sitting posture as her icon image because she had found that it was used by some users to make themselves look thinner.

In Information Retrieval (IR), Capra, et al. proposed the search assistance system, called "Search Guide" suing search trails [12]. Users can refer to other users' search trails (e.g., queries issued, results clicked, pages bookmarked) as an example of the search behaviour. It was found that the Search Guide can help users' complex search behaviour.

However, there is little work on the development and evaluation of tools to support the writing of profile text for SNS. In general, SNS profile text is much shorter than a story or scientific article. Therefore, we need a different approach from assisting users to maintain consistency in their writing or to write longer sentences efficiently [9, 10].

1.3. Research Aim

This study aims to help users identify what to write in their profile text. For this aim, we proposed a system to assist users in creating SNS profile text by searching relevant profile texts created by other SNS users. More specifically, we formulated the two research questions as follows.

RQ1 Do people generally find it difficult to write their SNS profile text, and if so why?

RQ2 Will the presentation of profile texts written by other users with similar interests help SNS users to write a longer profile text?

1.4. Paper Structure

The rest of the paper is structured as follows. In Section 2, the proposed writing assistant system is presented. In Section 3, we describe the user study to evaluate the effectiveness of the proposed system. In Section 4 presents the experimental results. In Section 5, we discuss the main findings and their implications on the SNS profile writing. And, finally, in Section 6, we conclude the paper with future work.

2. Proposed system

This section describes the major components of the proposed system: corpus, user interface, and back-end search system.

2.1. Corpus

The first step in our work was to build a corpus which was then indexed and searched by the proposed writing assistance system. Since the idea of observational learning indicates that users can benefit from the learning of prior examples created by other users in similar contexts, we decided to collect user profile texts from the accounts that deem to have some level of association with our study participants: students at one of the major universities in Japan.

First, we manually collected 46 Twitter lists which were created by the associated members of the university. Then we collected the accounts of the members of each collected list as well as the accounts who they follow. Finally, the data for each account, including the profile text, was automatically obtained via Twitter API. As a result, a total of 785,531 profile texts were collected.

Furthermore, we had two automated steps to remove profile texts for our research aim. First, as the collected accounts included accounts of organisations and bot accounts, we excluded them from the search if they contained any of the following words.

administrator / association / booking / bot / community / closed / event / info / official / sales / shop / tel

Second, it was necessary to exclude profile texts that were too short or did not contain enough information as observational learning. We decided to exclude profile texts with less than 40 characters from the search. In the end, 351,754 profile texts were included in the corpus.

2.2. User Interface

The UI of the proposed system is shown in Figure 2. The UI consists of three main areas: Edit Area, Search Result Area, and Keep Area. When a user edits a profile text on the form in Edit Area, the system retrieves relevant other users' profile text and displays them in the Search Result Area. The user can use the search results to refer to and rewrite his or her profile text. The search results were dynamically updated as the texts in Edit Area are changed. Therefore, we provided an option to "keep" profile texts you find helpful in the Keep Area. Finally,

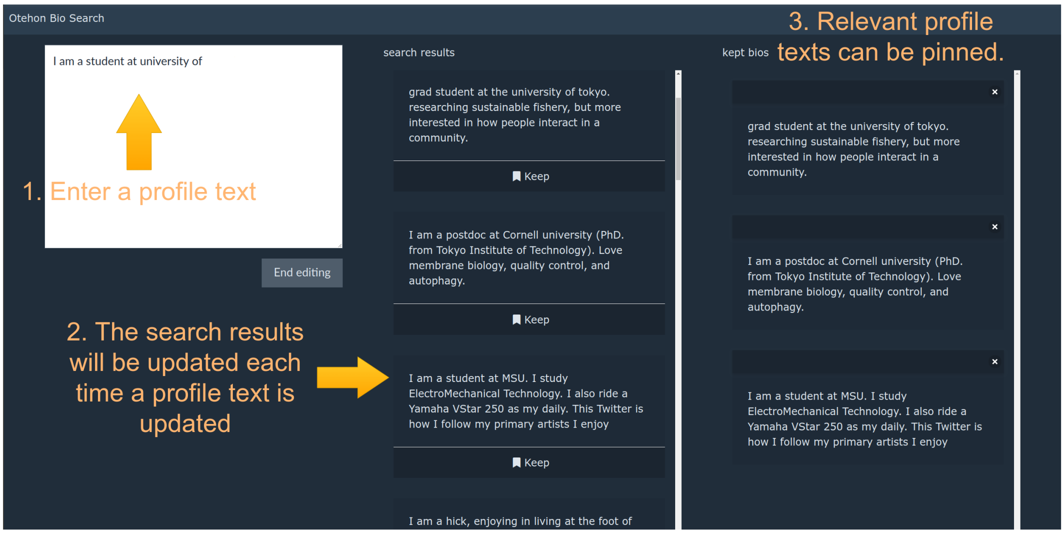

Figure 2: UI of the proposed system

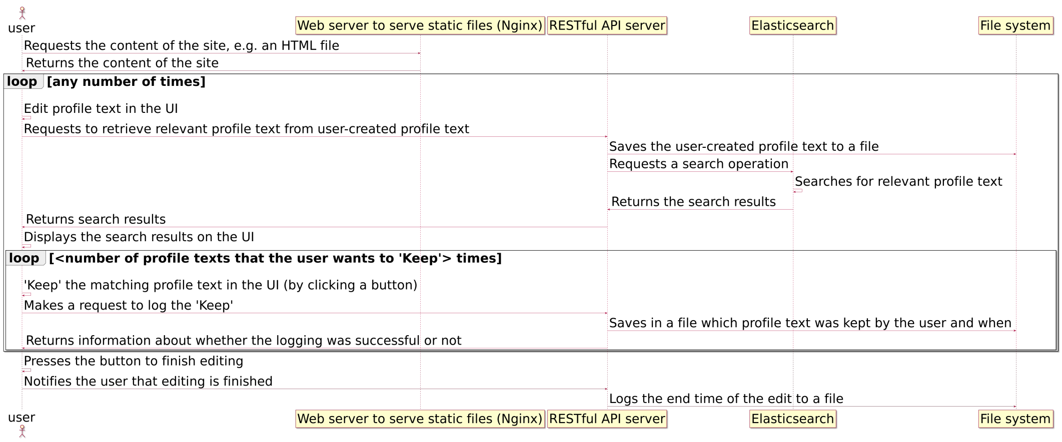

Figure 3: A sequence diagram of the proposed system

there was a button at the bottom of the Edit Area to indicate the completion of the writing task during the user study.

During the experiment, a controlled UI was also developed where the overall look was identical to the proposed UI. However, the controlled UI did not have the Search Result Area and Keep Area.

2.3. Back-End Search System

Figure 3 is a sequence diagram of the proposed system. The proposed system consists of three components: Static file server (Nginx), HTTP Server (RESTful API Server), and Elasticsearch. Each component was deployed as a Pod in an on-premises Kubernetes cluster. When a user accesses the URL of the proposed system front-end UI, the static file server first serves page contents such as HTML, CSS and JavaScript. After that, the interaction events on the page and the profile text created by the user are sent to the API server and stored in the server's file system. Finally, profile texts are retrieved from the corpus by Elasticsearch. We added some plugins to Elasticsearch to tokenize Japanese text and configured Elasticsearch to rank documents by Okapi BM25 [13].

Since we used a manually constructed non-standard corpus of SNS profile texts in our experiment, we tested the performance of the back-end search system to ensure that it can retrieve relevant texts. Twenty simulated queries were manually created by using partial profile texts, the top 200 documents were retrieved by the back-end search system, and their relevance was assessed with graded relevance by the authors. The MAP was 0.589 and nDCG@20 was 0.557. Although this was an informal system evaluation, the results were sufficiently promising for our research aim.

3. Experiments

To evaluate the effectiveness of the proposed system, a user study was conducted with 24 participants who used the interactive writing assistant system which indexed the custom collection of Twitter profile texts as described in 2.1. The user study was approved by the ethics committee of the Faculty of Library, Information and Media Science, Univerity of Tsukuba (No. 20-7). The experiment consisted of a pre-questionnaire, a profile text writing task and a post-questionnaire. Due to COVID-19, the call for participants and all tasks were conducted online.

3.1. Participants

Of 24 participants, 14 (58%) were female and 10 (42%) were male. All participants were undergraduate students in the age group between 18 and 22. Their academic background varied among Library and Information Science (13), Computer Science (2), Medicine (2), Social Sciences (2), Mathematics (1), Engineering (1), Disability Science (1), Education (1) and Media Science (1).

3.2. Profile Text Writing Task

Participants were asked to write a profile text in our user study. Participants were randomly assigned to one of two groups: Control and Experimental. Twelve participants in the Experimental group created a profile text using the proposed system, and twelve participants in the Control group created a profile text without using the suggestion function. They were also instructed to create a profile text under the following scenario.

“To increase online communication between students at the university, the university asked all students to create a Twitter account. What kind of profile text would you like to create for your SNS account for the campus life?”

Since our collection was focused on university-related user profiles, we designed the scenario as above to avoid zero-match results in the Search Results area during the profile writing task.

Table 1
Answers to “Please answer only if you have used SNS before: Have you ever had trouble writing your profile text?” (5-point scale from “1. Never” to “5. Every time I cannot write well”, 24 participants)

Answer	Number of participants
5 (Every time I cannot write well)	1 (4%)
4	10 (42%)
3	4 (17%)
2	7 (29%)
1 (Never)	2 (8%)

4. Results

A total of 24 profile texts were generated in the experiment and used for analyses.

4.1. Profile Writing Experience

To answer RQ1, we investigated participants’ profile writing experience. We prepared a questionnaire that asked: "Have you ever had trouble writing in the self-description text of your SNS profile? The breakdown of the answers (on a five-point scale from “1. Never” to “5. Every time I cannot write well”) is shown in Table. 1. Only 8% of the participants answered “1. Never”, indicating that having a difficult experience in writing their SNS profile text is common. The average of the answers was 3.04 and the standard deviation was 1.12.

A follow-up question was asked only for those participants who had experienced difficulties (Answer 2 to 5) in writing their profile text: “What is the reason for this?” The options of answers and the results are shown in Table 2. The most common answer was “I could think of many things I could write about, but I didn’t know what to write about in particular” (9 of 18 participants answered). The second most common answer was “I didn’t feel I had anything to write about” and “I wanted to remain anonymous, so I had to be careful not to include any personal information” (7 of 18). Again, these responses support our motivation that many SNS can benefit from writing assistant tools to enrich their profile texts, although some users intentionally provided limited information to keep their privacy.

4.2. Profile Texts

To answer the research question RQ2, we compare the distribution of the number of characters in the profile text created between the two groups. Firstly, the number of characters in the profile text created by each group in descending order is compared in Figure 4. As can be seen, the experimental group tended to have more characters

Table 2
Please answer this question only if you answered in the previous question that you have had trouble writing your profile text. What was the reason? (multiple answers possible, 18 respondents)

Selected reason	Number of participants
"I could think of lots of things I could write about, but didn't know what to write about in particular"	9 (50%)
"I didn't feel I had anything to write about"	7 (39%)
"I wanted to remain anonymous, so I had to be careful not to include any personal information"	7 (39%)
"I was concerned that I might come across as self-conscious if I wrote long sentences"	6 (33%)
"I thought it would be strange if I wrote something different from other users around me"	2 (11%)
Others (e.g. "I didn't know how to introduce myself")	2 (11%)

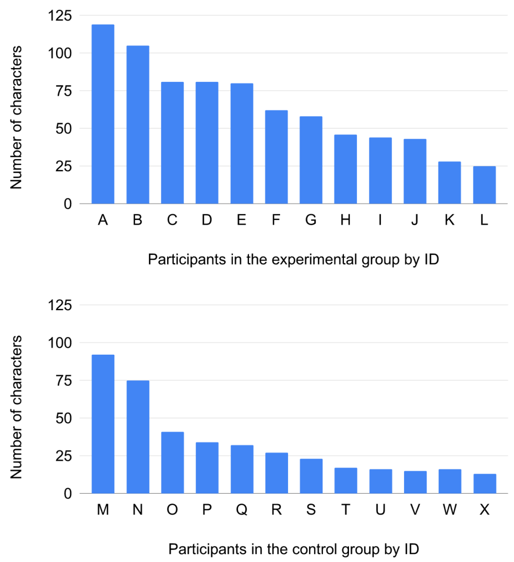

Figure 4: Distribution of the number of characters in the created profile text (N=12)

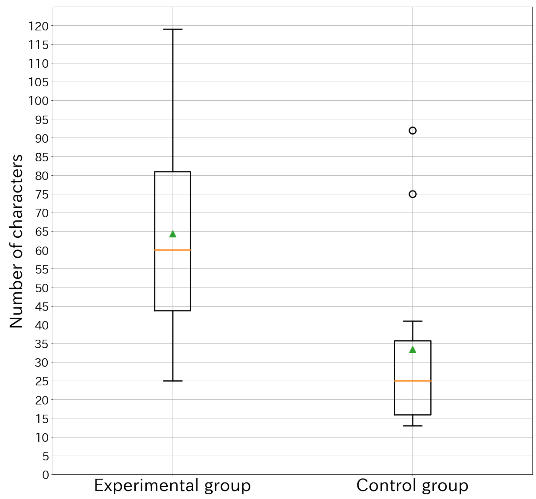

Figure 5: Distribution of the number of characters in the created profile text - comparison by boxplot (N=12)

in their profile texts.

Next, we compared it using the boxplot and the results are shown in Figure 5. It can be seen that the quartiles and the minimum and maximum values are larger in the experimental group. In addition, the value of the quartile range is relatively larger in the experimental group. This indicates that the effect of increasing the number of characters by using the proposed system varied across participant. The descriptive statistics values are summarized in Table 3.

Furthermore, a statistical test was carried out to see if there was a statistical significance in the distribution of the number of characters between the two groups. The Shapiro-Wilk test was used to confirm that the distribution did not follow a normal distribution for both groups, and thus, the Mann-Whitney U test was used as the testing method. It is a two-tailed test and the null hypothesis is rejected at 5% level of significance. Since the sample size in this study is small (12 participants for each groups), we thought it necessary to also focus on the effect size as well as the p-value [14]. The effect size r for the Mann-Whitney U test can be calculated from the test statistic U and the sample size and satisfies $-1.00 \leq r \leq 1.00$ [14]. The coin package [15] (1.4-1) on the CRAN package manager for R language was used to compute p-value and effect size. The results of the Mann-Whitney U test are shown in Table 4. From $p < 0.0.5$, there is a significant difference between the two distributions. In addition, as r-value coincides with the one for Pearson's correlation coefficient [14], it can be considered as a moderate effect size $(0.36 \leq r \leq 0.67)$ [16].

Additionally, we analysed the time participants spent during the profile writing task, and Figure 6 shows the result. It was found that the experimental group took 164.5 seconds longer than the control group on the median

Table 3
The descriptive statistics values of the distribution of the number of characters

	Min.	1st Quart.	Central	3rd Quart.	Max.	Quart. Range	Mean	SD
experimental system	25.00	43.75	60.00	81.00	119.00	37.25	64.33	29.48
control system	13.00	16.00	25.00	35.75	92.00	19.75	33.42	25.23

Table 4
Results of a Mann-Whitney U-test on the distribution of the number of characters in the profile texts for both groups

p-value	Z statistics	Effect Size r	Effect Size [16]
0.04307	2.7725	0.5659342	moderate

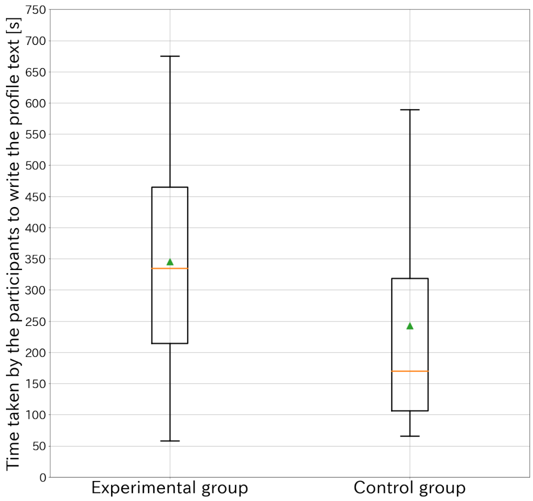

Figure 6: Time taken by the participants to write their own profile text [s] (N=12)

comparison, suggesting that participants in the experimental group spent more time on profile writing.

4.3. Items in Profile Texts

We asked the following questions in the post questionnaire: "What did you have in mind when you started writing your profile text?" (A) and "What was written in the final profile text?" (B). In both questions, participants selected their answers from the same list of possible answers, including "Affiliation" and "Hobbies". We counted the number of increase between what participants "Intended to Write" (A) and what they "Actually Wrote" (B). The result is summarized in Table 5. The most common answer was "Others" ("Research Interests", "Past Affiliations", etc.), followed by "URLs to their social networking sites or blog", and "Hobbies and Interests".

Table 5
Breakdown of "Items that participants did not intend to write but actually wrote" (multiple answers) in the experimental group.

Item name	Number of participants
Other (Free answer. "Research Interests", "Previous Affiliation", "Greetings", "Awards Received", "Hometown")	5
URLs (to other social networking sites or blog)	4
Hobbies and Interests	3
Age	2
Name/Nickname/What you are called	1

While RQ2 suggests that the proposed system allowed participants to write longer texts in the SNS profile, the answers to this questionnaire give more details about how those longer texts enriched the profile texts that were otherwise shorter and potentially less diverse.

5. Discussion

This section highlights the main findings from our study and discusses their implication on supporting profile writing and the limitation of the study.

First of all, we discuss the findings on RQ1 "Do people generally find it difficult to write their SNS profile text, and if so why?" The outcome of the questionnaire suggests that 44% of participants found some level of difficulty in writing profile texts. This is a large proportion given the scale of SNS user populations. The cause of difficulty varied from the lack of ideas to uncertainty of self-disclosure and self-impression, to a concern of privacy. This finding supports our intuition on the need for assistance in writing profile texts in SNS.

RQ2 was "Will the presentation of profile texts written by other users with similar interests help SNS users to write a longer profile text?" Participants managed to write significantly longer profile texts with the proposed system where existing profile texts were dynamically retrieved and presented to the writing UI. The follow-up questionnaire showed that participants obtained the ideas of adding their interests and relevant URLs to the profile texts in the proposed system.

Therefore, providing a learning opportunity based on existing users with similar interests seems to be a promising method to help SNS users to enrich their profile texts. Unlike the assistance system in academic writing or creative writing[9, 10], the profile writing seems to benefit from the prior examples of "similar account" since the main aim of SNS is to connect with people with similar interests. Given that online communities are diverse, our approach of retrieving profile texts from similar accounts seems to have an advantage over a fixed list of items to include in the profile.

As for the limitation of our work, first, participants of our user study were recruited from a single university although their academic background varied. Therefore, a similar effect might not appear in other SNS user populations (e.g., different age groups or occupations). Second, this study did not verify whether the profile text created by the proposed system is indeed effective at expanding social network. Third, as pointed out in previous research[17], user privacy also needs to be considered in supporting the creation of user-generated content, including profile text. Finally, this study did not investigate whether the ranking algorithm of the back-end search system affected the user experience.

6. Conclusion and Future Work

Based on our preliminary observation that the average length of SNS profile texts was not nearly as long as it could be, we developed a search-based profile writing support tool. The basic idea of our proposed system was to provide users with potential ideas about what to write and how to express by offering profile texts written by other users with similar interests. A user study was carried out with 24 participants to evaluate the effectiveness of the proposed system. The experimental results show that the proposed system can enhance the profile writing task by allowing users to include more items in the profile texts leading to longer and richer contents. The results of questionnaires also indicate that one of the reasons for short profile texts is due to the difficulty in writing profile texts, supporting the motivation of our work, while some users preferred to be anonymous with limited profile content. Nevertheless, this work suggests that the proposed system will be helpful for those who would like to effectively extend their social network by richer profile contents.

This paper demonstrated that search-based assistance was effective for SNS profile writing of a particular university student group. Future work should investigate the effectiveness with other university student groups as well as other population groups (e.g., professionals). Further work is also desirable to examine whether the profiles written by the proposed system can lead to a better extension of social networks than existing profiles or profiles written by other methods. Furthermore, search-based observational learning seems versatile enough to apply to other domains, and thus, the effectiveness of the proposed system in other writing tasks such as product review writing should also be an interesting research direction. Finally, how to achieve a good balance between privacy and effective profile is yet another important research question to be addressed in future work. From a system design perspective, how to integrate the assistance function to the operational SNS needs to be examined [18].

References

[1] P. T. Metaxas, E. Mustafaraj, K. Wong, L. Zeng, M. O'Keefe, S. Finn, Do Retweets indicate Interest, Trust, Agreement? (Extended Abstract), CoRR abs/1411.3 (2014). URL: http://arxiv.org/abs/1411.3555. arXiv:1411.3555.

[2] N. Ellison, R. Heino, J. Gibbs, Managing Impressions Online: Self-Presentation Processes in the Online Dating Environment, Journal of Computer-Mediated Communication 11 (2017) 415–441. URL: https://doi.org/10.1111/j.1083-6101.2006.00020.x. doi:10.1111/j.1083-6101.2006.00020.x.

[3] S. T. Tong, E. F. Corriero, K. A. Wibowo, T. W. Makki, R. B. Slatcher, Self-presentation and impressions of personality through text-based online dating profiles: A lens model analysis, New Media & Society 22 (2020) 875–895.

[4] X. Ma, J. T. Hancock, K. Lim Mingjie, M. Naaman, Self-disclosure and perceived trustworthiness of airbnb host profiles, in: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW '17, Association for Computing Machinery, New York, NY, USA, 2017, p. 2397–2409. URL: https://doi.org/10.1145/2998181.2998269. doi:10.1145/2998181.2998269.

[5] S. Counts, K. Stecher, Self-presentation of personality during online profile creation, in: Conference: Proceedings of the Third International Conference on Weblogs and Social Media, ICWSM 2009, 2009.

[6] H. Krasnova, S. Spiekermann, K. Koroleva, T. Hildebrand, Online social networks: why we disclose, Journal of Information Technology 25 (2010) 109–125. URL: https://doi.org/10.1057/jit.2010.6. doi:10.1057/jit.2010.6.

[7] F. Stutzman, J. Kramer-Duffield, Friends only: Examining a privacy-enhancing behavior in facebook, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '10, Association for Computing Machinery, New York, NY, USA, 2010, p. 1553–1562. URL: https://doi.org/10.1145/1753326.1753559. doi:10.1145/1753326.1753559.

[8] A. Bandura, Observational Learning, 2 ed., Macmillan Reference USA, New York, NY, 2004, pp. 482–484. URL: https://link.gale.com/apps/doc/CX3407100173/GVRL?u=cuny_hunter&sid=GVRL&xid=369c6e06, topic overview.

[9] M. Roemmele, A. S. Gordon, Creative help: A story writing assistant, in: International Conference on Interactive Digital Storytelling, Springer, 2015, pp. 81–92.

[10] T. Kinnunen, H. Leisma, M. Machunik, T. Kakkonen, J.-L. LeBrun, SWAN - scientific writing AssistaNt. a tool for helping scholars to write reader-friendly manuscripts, in: Proceedings of the Demonstrations at the 13th Conference of the European Chapter of the Association for Computational Linguistics, Association for Computational Linguistics, Avignon, France, 2012, pp. 20–24. URL: https://www.aclweb.org/anthology/E12-2005.

[11] M.-H. Chen, S.-T. Huang, H.-T. Hsieh, T.-H. Kao, J. S. Chang, FLOW: A first-language-oriented writing assistant system, in: Proceedings of the ACL 2012 System Demonstrations, Association for Computational Linguistics, Jeju Island, Korea, 2012, pp. 157–162. URL: https://www.aclweb.org/anthology/P12-3027.

[12] R. Capra, J. Arguello, A. Crescenzi, E. Vardell, Differences in the use of search assistance for tasks of varying complexity, in: Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '15, Association for Computing Machinery, New York, NY, USA, 2015, p. 23–32. URL: https://doi.org/10.1145/2766462.2767741. doi:10.1145/2766462.2767741.

[13] S. Robertson, S. Walker, S. Jones, M. M. Hancock-Beaulieu, M. Gatford, Okapi at trec-3, in: Overview of the Third Text REtrieval Conference (TREC-3), Gaithersburg, MD: NIST, 1995, pp. 109–126. URL: https://www.microsoft.com/en-us/research/publication/okapi-at-trec-3/.

[14] M. Tomczak, E. Tomczak, The need to report effect size estimates revisited. an overview of some recommended measures of effect size 21 (2014) 19–25.

[15] H. Torsten, H. Kurt, M. Van de Wiel, Z. Achim, Implementing a class of permutation tests: The coin package, Journal of Statistical Software 28 (2008). doi:10.18637/jss.v028.i08.

[16] R. Taylor, Interpretation of the correlation coefficient: A basic review, Journal of Diagnostic Medical Sonography 6 (1990) 35 – 39.

[17] Y. Fang, A. Godavarthy, H. Lu, A utility maximization framework for privacy preservation of user generated content, in: Proceedings of the 2016 ACM International Conference on the Theory of Information Retrieval, ICTIR '16, Association for Computing Machinery, New York, NY, USA, 2016, p. 281–290. URL: https://doi.org/10.1145/2970398.2970417. doi:10.1145/2970398.2970417.

[18] G. Dworman, S. Rosenbaum, Helping users to use help: Improving interaction with help systems, in: CHI '04 Extended Abstracts on Human Factors in Computing Systems, CHI EA '04, 2004, p. 1717–1718. URL: https://doi.org/10.1145/985921.986198. doi:10.1145/985921.986198.