Evaluation of a Chat Interface for Diary Studies in Information Behavior Research

Konstantin Kulik
Media Informatics Group
University of Regensburg,
Germany
kulik.konstantin@gmx.de

Michael Achmann
Media Informatics Group
University of Regensburg,
Germany
michael.achmann@ur.de

Christian Wolff
Media Informatics Group
University of Regensburg,
Germany
Christian.wolff@ur.de

ABSTRACT
We report about our work in progress evaluation of a Telegram-based chat-interface for collecting diary entries in the field of information behavior research and compare the results to web-form based entries. We expect participants to complement their chat entries by multimedia files and engage more openly.

KEYWORDS
diary studies; chat bot; ethnographic methods; information behavior; research methods.

INTRODUCTION
With the current pandemic limiting our choice for ethnographic studies in information behavior, we were looking for remote observational methods to be used in an ongoing project. Investigating the activities in the scholarly research process, diary studies quickly attracted our attention, as they are staple methods for research in our field (e.g. Case, 2012, p. 222 ff. Sohn, Li, Griswold, & Hollan, 2008) and offer the possibility to gather a multitude of data, like e.g. "personal testimonies", "everyday events", "longitudinal or chronological details" and more (Hyers, 2018). The possible study designs are at least as diverse as the data they may collect; thus, we chose to use a chat-interface to elicit daily activities of students, to understand the problems they encounter, the progress they make, and the role of software in their daily work. We present our work in progress evaluation of a chat interface for collecting diary entries, asking the two main questions: What data does it collect and how is it different from plain web forms?

RELATED WORK
Mobile phones have been used in the past to conduct diary studies. Palen & Salzman (2002) presented a voice-mail diary study, praising the method for the opportunity to study activities naturalistically and with minimal intrusion, especially when using mobile phones. Supplementing voicemail Brandt, Weis, & Klemmer (2007) using text and picture messages to capture snippets of information in situ which were followed up in a web form eliciting more information for each snippet. Just little later Sohn et. al. (2008) conducted an influential study of mobile information needs, using text messages to capture diary entries in situ. Their study used several reminders and they also captured snippets via text messages, followed up by a daily web survey to add additional data to their text message snippets.

More recently studies by Kaufmann & Pfeil (2020) and Kim et al. (2019) used contemporary instant messengers. Kaufmann and Pfeil utilized WhatsApp to perform a mobile instant messaging interview to gather qualitative data. This method showed several advantages like not being intrusive in everyday life, encouraging informal interaction which stimulates open communication, and the data being more diverse as the participants could send pictures and other media. Kim et al. went another direction. They investigated whether text-based chatbots are a better alternative towards web-based questionnaires. They showed that participants who answered questions from the chatbot gave more diverse and less satisfactory answers.

All in all (mobile) phones have proofed useful to capture diary entries. In the past, however, text snippets needed to be complemented by surveys and structuring voice diary study was complex. Modern messaging applications answer some of the shortcomings in the first mobile diary studies. The literature proofs their use for interview studies and positive effects as an alternative interface for questionnaires. Thus, we investigate how mobile diary studies may be improved using a chat interface using instant messaging.

METHOD
The Chat Interface
We developed a chat bot using the Telegram platform. While called "Bot" by the Telegram documentation, we decided to use the term "chat interface" to prevent any misinterpretations: Our implementation is not "smart", thus we are not making use of any Natural Language Processing, rather platform buttons or commands are used. As the Telegram app is intrinsically linked to mobile devices, multimedia messages, like speech, video and image messages are permitted as they are part of the "typical" chat experience, offered by many contemporary chat apps including e.g. WhatsApp.

84th Annual Meeting of the Association for Information Science & Technology | Oct. 29 – Nov. 3, 2021 | Salt Lake City, UT.

The bot is implemented using the python-telegram-bot package (at https://github.com/python-telegram-bot/python-telegram-bot). The python script connects to several Google Spreadsheets storing user input and user information. Once a user has signed up, their incoming messages are recognized by the users' telegram-account-id. This id is used to link information across the tables of our semi-structured diary study. Messages that contain media files will be forwarded to the research directors personal telegram account as the official Telegram API only allows to download files which are up to 20 MB large. With the forwarding method we secure to collect every media file that is sent to the bot, as video files, for example, quickly exceed the limit.

The Form Interface

We decided to evaluate the chat interface against an online form using Google Forms. We used the same questions guiding the diary when creating a form with free-text input fields and an additional file upload field for each text field. File upload was limited to 10 files per question for technical reasons. (This is the maximum number of files Google is allowing to upload per file upload field.) Submitted forms are saved in a spreadsheet format on Google Drive, thus they may be exported easily for further data analysis.

The Study

To evaluate the chat interface, we designed a diary study to elicit daily activities and problems during the scholarly research process. The study consisted of 4 short open-ended questions (see Table 1) to be answered by the participants during the period of two weeks once a day. Two types of interfaces were offered to create a diary entry: A google form interface or a Telegram chat interface. Both allowed writing entries in a text field or sending / uploading media. The latter option was not emphasized, however all interfaces offer sending audio messages, images, and other media, allowing both, to create an entry by e.g., speaking with the device, or complementing the entry by artefacts. In our within-subject study each participant was asked to use the one interface for a continuous week, then the other. To control for unwanted effects due to the change between interfaces, the participants were randomly split into two groups. Group A started using the chat-interface, group B the web-forms. Following the week-long period for each interface-type, participants were asked to answer the System Usability Scale (Bangor, Kortum, & Miller, 2008; Brooke, 1996) and the NASA-TLX (Hart & Staveland, 1988) to assess the perceived usability and perceived workload of the web and chat applications. These questionnaires were embedded into an independent, self-developed website.

Q1	What have you accomplished today regarding your university work?
Q2	Are you happy with all you've accomplished today? Please elaborate.
Q3	Is there anything that might have worked better today? If so, what might have helped?
Q4	Would you like to tell us anything else?

Table 1. Diary Study Questions

Participants were reminded daily at 5 o'clock using a Telegram message to submit a diary entry. Chat messages were triggered for both interfaces to control for any effects another medium for notifying the user might have had. Following the study, participants are asked to partake in a follow-up interview. They are asked to elaborate about unclear entries and about their personal preference between the two capturing techniques.

FUTURE WORK

At the time of writing 51 participants signed up for the study which has not yet been concluded. The collected data will be analyzed quantitatively and qualitatively: We are interested in the differences between collected responses by the length of the data, the number of responses and type of responses: How often do participants use the multimedia options? Further it may be interesting to see whether emoticons are used and how much time passes between reminder and submitting the entry. Speech messages, another stock feature of modern chat applications, might blend the diary study with a traditional interview. From the qualitative point of view, we are planning to code the responses using a grounded theory approach (Corbin & Strauss, 2008). These processes may be complemented by frequency analysis and topic modelling techniques.

DISCUSSION

If the chat interface's responses provide a more thorough look into the work and problems of our subjects, we are confident to use and recommend this technique for further research using diary studies. The implementation of a chat interface is more time-consuming, thus the additional information captured needs to justify the extra work. While comparing different methods for capturing information in diary studies we see two potential shortcomings in our design: We need to evaluate whether our design measured information behavior once data collection concluded and, while exploring the technique as a COVID-compliant approach to ethnographic studies, we are not evaluating the diary study itself against in-situ or other ethnographic approaches.

REFERENCES

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human–Computer Interaction*, *24*(6), 574–594.

Brandt, J., Weiss, N., & Klemmer, S. R. (2007, April 28). txt 4 l8r: lowering the burden for diary studies under mobile conditions. *CHI '07 Extended Abstracts on Human Factors in Computing Systems*, 2303–2308. Presented at the San Jose, CA, USA. New York, NY, USA: Association for Computing Machinery.

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. Jordan, B. Thomas, B. Weerdmeester, & I. L. McClelland (Eds.), *Usability Evaluation In Industry* (pp. 207–212). CRC Press.

Case, D. O. (2012). *Looking for Information: A Survey of Research on Information Seeking, Needs and Behavior*. Emerald Group Publishing.

Corbin, J. M., & Strauss, A. L. (2008). *Basics of qualitative research: techniques and procedures for developing grounded theory*. Sage Publications, Inc.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in Psychology* (Vol. 52, pp. 139–183). North-Holland.

Hyers, L. (2018). *Diary Methods: Understanding Qualitative Research*. New York: Oxford University Press.

Kaufmann, K., & Peil, C. (2020). The mobile instant messaging interview (MIMI): Using WhatsApp to enhance self-reporting and explore media usage in situ. *Mobile Media & Communication*, *8*(2), 229–246.

Kim, S., Lee, J., & Gweon, G. (2019). Comparing data from chatbot and web surveys. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. Presented at the the 2019 CHI Conference, Glasgow, Scotland Uk. doi:10.1145/3290605.3300316

Palen, L., & Salzman, M. (2002, November 16). Voice-mail diary studies for naturalistic data capture under mobile conditions. *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work*, 87–95. Presented at the New Orleans, Louisiana, USA. New York, NY, USA: Association for Computing Machinery.

Sohn, T., Li, K. A., Griswold, W. G., & Hollan, J. D. (2008). A Diary Study of Mobile Information Needs. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 433–442. Presented at the Florence, Italy. New York, NY, USA: ACM.