

Crowdsourcing Long COVID Sweden (CiLC-S): Exploring Digital Methods and Voluntary Health Data for Research and Response in Social Crises

Veronica Johansson <https://orcid.org/0000-0001-5523-0282>

Swedish School of Library and Information Science (SSLIS), The Academy for Libraries, Information, IT and Pedagogy, University of Borås, Borås, Sweden

veronica.johansson@hb.se

Linda Wänström

Department of Computer and Information Science, The Division of Statistics and Machine Learning, Linköping University, Linköping, Sweden

linda.wanstrom@liu.se

Kim Ramme

Moment Children and Youth Medical Centre, Vällingby, Sweden

Kim@Rammes.se

Sakari Kallio

School of Bioscience, University of Skövde, Skövde, Sweden

sakari.kallio@his.se

Johanna Nilsson Swartwe

School of Bioscience, University of Skövde, Skövde, Sweden

johanna.swartwe@gmail.com

Kayvan Yousefi Mojir <https://orcid.org/0000-0002-9781-2106>

Swedish School of Library and Information Science (SSLIS), The Academy for Libraries, Information, IT and Pedagogy, University of Borås, Borås, Sweden

kayvan.yousefi.mojir@hb.se

Keywords

COVID-19, data, information needs, information strategies

EXTENDED ABSTRACT

Topic

We describe here the emergent, multi-disciplinary research project CiLC-S – Crowdsourcing Long COVID Sweden. The information science related part of the project aims to explore i) information needs and strategies of persons afflicted by long term health problems after a COVID-19 infection, and ii) alternative digital methods for inclusive, safe and efficient data collection for research and social response in the COVID-19 pandemic. The purpose of this presentation is to introduce the project and share some early ethical and methodological experiences as a foundation for discussions of current and future possibilities for information science research to respond to individual and societal information needs in the COVID-19 pandemic and similar large-scale social crises through innovative research questions and data collection methods.

Value

CiLC-S is a collaborative research project, constructed around a problem with overlapping interests of established research and social actors. In the project, researchers from information science, statistics, medicine, biology, and cognitive neuroscience collaborate across disciplines and with the non-profit patient organisation the Swedish COVID Society around research questions related to the group of so-called 'long-haulers' (persons with health problems lasting more than three months after a COVID-19 infection) in Sweden. There are three main expected values and results from the project. The first is to contribute to increased knowledge of long COVID (or post COVID) in Sweden concerning symptoms and other characteristics of the illness among afflicted individuals, with associated consequences on individual and societal levels. The second is to generate insights into how an improved information infrastructure for the information needs of afflicted individuals, the civil society and public institutions in the corona pandemic and similar large-scale social crises can be supported. The third is to develop and adapt crowdsourcing based methods for inclusive, safe and efficient data collection and analysis in time restricted, large-scale crises for research and social response.

Methodological challenges and lessons thus far include: dealing with institutional barriers to inter-university and inter-organisational collaboration on potentially sensitive personal data and data collection instruments (survey platform); seeking measures to control and compensate for authenticity and representativity uncertainty in open, anonymous, questionnaires; finding means to make possible longitudinal, follow-up, surveys and similar data collection among the same anonymous individuals in the cohort; transforming commonly qualitative information needs and practices interview questions into quantitative survey adapted counterparts; planning for ensuring anonymity for answers to open questions in the extensive survey with presumed very large response rate (large volumes of free text data); and finding ways to make such data openly accessible and useful to others both quickly and in the long term.

Research outline

In reaction to slow, inefficient and potentially biased research on patients with formal diagnoses in hospital records, the value of patient inclusive, participatory forms of medical research projects seem to increase. In these types of projects, the directly concerned subjects (patients) are represented, involved in decision making and/or leading in data collection and problem formulation (e.g., von der Lehr, 2021; Iwata et al., 2019; Hosio et al., 2018). This is believed to lead to faster, more efficient and representative methods for data collection (Dhruva et al., 2020), in particular in crisis like situations such as the COVID-19 pandemic (Davis et al. 2020; Kent, 2020). Compared to the social science (e.g., Marres, 2017; Hargittai, 2021), research in health care and medicine have been late to embrace internet adapted digital methods although studies indicate positive results. Participant recruitment through Facebook shows good representativity in general as well as for traditionally difficult to reach demographics

(Whitaker, Stevelink & Fear, 2017); social media, apps, and smartphones show good results for clinical (Rosa et al., 2015) as well as medical (Topolovec-Vranic & Natarajan, 2016) research; as do projects with patient driven or citizen science related design (t.ex. Collins et al., 2018; Iwata et al., 2019).

Further, it is assumed that patients will pose increasingly high demands on reception of, insight into, and influence over, relevant data collection and formulation of research questions (Dhruva et al., 2020; Hosio et al., 2018). Data collection that goes further in this direction relies on e.g. crowdsourcing methods and individuals' and patients' voluntary sharing of health data –from both hospital records and personal health technology (smart watches, health tracking applications) – with researchers and each other (e.g., Desai, Kuderer & Lyman, 2020; Fliesler, 2020; Hosio et al., 2018; Hosio et al., 2020; Kent, 2020). However, in these contexts other types of problems arise, not least concerning the validity, authenticity and representativity of data. Research aiming to develop methods for evaluation and quality assurance of such new data sources is underway (e.g. Lane, Armin & Gordon, 2015; Marres, 2017; Checco, Bates & Demartini, 2020; Kalimeri et al., 2020; Hargittai, 2021), but the territory is new and much remains to be developed and evaluated.

In yet another perspective, the COVID-19 pandemic and the situation of long haulers also raise questions concerning the Swedish society's and other nations' abilities and preparedness for handling large-scale health crises from a societal information perspective. Crisis situations in general lead to individuals' development of new and sometimes unforeseeable information needs and strategies with similarly unforeseeable consequences (e.g. Hagar, 2010), which further underlines the importance of information science studies on the COVID-19 pandemic (Lloyd & Hicks, 2020; Xie et al., 2020). Not least citizens' usage of social media leads to specific opportunities as well as risks. Research shows, e.g., that governments and public agencies need to include social media in communication with and warnings to the public (e.g., Brynielsson et al., 2018) but also that the public and individuals are vulnerable in crisis situations such as the corona pandemic with obvious risks of psychological ill-health and exposure to disinformation (Su et al., 2021; c.f. also Shklovski, Palen, & Sutton, 2008).

There are, thus, numerous information related research questions in the wake of the COVID-19 pandemic that are deeply connected to medical correspondences, and equally urgent to study. TheCiLC-S project comprises, against this background, three overarching research questions that correspond to one work package (WP):

- WP1: What demographics, symptoms and health and work related experiences characterise persons in Sweden who describe themselves as suffering from health problems for more than three months after a COVID-19 infection?
- WP2: What information needs and strategies can be identified among Swedish 'long haulers'?

- WP3: How can crowdsourcing based digital methods be adapted for inclusive, safe and efficient data collection for research and social response on the COVID-19 pandemic and similar large-scale social crises?

Empirical data is primarily planned to be collected through open, national questionnaires with both closed and open questions, distributed to members of Swedish Facebook groups for COVID-19 long haulers, through the Swedish COVID Society's e-mail list, and the research project's website. The anticipated response rate is around 2000 answers. Another study plans for a smaller set of qualitative, in depth focus group interviews with moderators of the same Facebook groups.

References

- Checco, A., Bates, J. & Demartini, G. (2020). Adversarial attacks on crowdsourcing quality control. *Journal of Artificial Intelligence Research*, 67. <https://www.jair.org/index.php/jair/article/view/11332>
- Collins, S. E. et al. (2018). Community-based participatory research (CBPR): Towards equitable involvement of community in psychology research. *The American psychologist*, 73(7), 884–898.
- Davis, H. E. et al. (2020). Characterizing long COVID in an international cohort: 7 months of symptoms and their impact. *Medrxiv*, December 27, December, 2020.
- Desai, A., Kuderer, N. M. & Lyman, G. H. (2020). Crowdsourcing in crisis: Rising to the occasion. *JCO Clinical Cancer Informatics*, 4, 551—554.
- Dhruva, S. et al. (2020). New federal rules pave the way for patient-driven health information exchange and real-world evidence on COVID-19 surveillance and treatment. (Blog post). *Health Affairs*, May 10, 2020. <https://www.healthaffairs.org/doi/10.1377/hblog20200506.368396/full/>
- Fliesler, N. (2020). Crowdsourcing COVID-19: New website tracks user-reported symptoms in real time. *News & Research. Harvard Medical School*. April 9, 2020. <https://hms.harvard.edu/news/crowdsourcing-covid-19>
- Hagar, C. (2010). Whom do you trust? Information seeking during the U.K. foot and mouth crisis. *Library & Archival Security*, 23(1), 3–18.
- Hargittai, E. (Ed.) (2021). *Research Exposed: How Empirical Social Science Gets Done in the Digital Age*. New York: Columbia University Press.
- Hosio, S., Karppinen, J., van Berkel, N., Oppenlaender, J. & Goncalves, J. (2020). Mobile decision support and data provisioning for low back pain. *IEEE Computer*, 51(8).
- Hosio, S., Karppinen, J., Takala, E. P., Takatalo, J., Goncalves, J., van Berkel, N., Konomi, S., Kostakos, V. (2018). Crowdsourcing treatments for low back pain. *Proc. ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2018)*.

- Iwata, A. J. et al. (2019). Flexible model for patient engagement: Achieving quality outcomes and building a research agenda for head and neck cancer. *Head & neck*, 41(4), 1087–1093.
- Kent, J. (2020). Supporting secure data sharing, patient privacy during COVID-19. *Health IT Analytics*, November 17, 2020.
- Lane, T. S., Armin, J., & Gordon, J. S. (2015). Online recruitment methods for web-based and mobile health studies: A review of the literature. *Journal of Medical Internet rEsearch*, 17(7), e183. <https://doi.org/10.2196/jmir.4359>
- Lloyd, A., & Hicks, A. (2020). Risk and resilience in radically redefined information environments; Information practices during the COVID-19 pandemic. *Proc Assoc Inf Sci Technol*. 57: e336. <https://doi.org/10.1002/pr2.336>
- Marres, N. (2017). *Digital Sociology: The Reinvention of Social Research*. Malden, MA: Polity.
- Rosa, C., Campbell, A. N. C., Miele, G. M., Brunner, M. & Winstanley, E. L. (2015). Using e-technologies in clinical trials. *Contemporary Clinical Trials*, 45, 41–54.
- Shklovski, I., Palen, L., & Sutton, J. (2008). Finding community through information and communication technology in disaster events. In *Proceedings of the 2008 ACM conference on computer supported cooperative work*, 127–136, New York, NY: ACM Press.
- Su, Z. et al. (2021). Mental health consequences of COVID-19 media coverage: The need for effective crisis communication practices. *Globalization and health*, 17(1), 4–4.
- Topolovec-Vranic, J. & Natarajan, K. (2016). The use of social media in recruitment for medical research studies: A scoping review. *J Med Internet Res*, 18(11).
- Whitaker, C., Stevelink, S. & Fear, N. (2017). The use of Facebook in recruiting participants for health research purposes: A systematic review. *J Med Internet Res*, 19(8). doi: 10.2196/jmir.7071
- Xie, B., He, D., Mercer, T., Wang, Y., Wu, D., Fleischmann, K. R., ... Lee, M. K. (2020). Global health crises are also information crises: A call to action. *Journal of the Association for Information Science and Technology*. <https://doi.org/10.1002/asi.24357>

Suggested Citation

Johansson, V., Wänström, L., Ramme, K., Kallio, S., Nilsson Swartswe, J. & Yousefi Mojir, K. (2021). Crowdsourcing Long COVID Sweden (CiLC-S): Exploring digital methods and voluntary health data for research and response in social crises. Trends in Information Science – The ASIS&ST European Chapter Research Series. 09-11 June 2021. Virtual conference.

© 2021. This work is licensed under a [CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/).