Exploring research data management practices among scholars that conduct data-intensive social sciences and humanities research

Anna Sendra
Tampere University
anna.sendratoset@tuni.fi
https://orcid.org/0000-0001-7069-1378

Elina Late
Tampere University
elina.late@tuni.fi
https://orcid.org/0000-0002-3232-1365

Sanna Kumpulainen
Tampere University
sanna.kumpulainen@tuni.fi
https://orcid.org/0000-0002-7016-257X

Introduction

The data deluge (Poole, 2016) present in contemporary societies has increased the interest in data management, also in academia (Nahotko et al., 2023). However, not all disciplines have been engaging in research data management practices equally, with fields pertaining to science, technology and medicine more used to exercising this type of activities (Borgman, 2007; Khan et al., 2023). Moreover, research data services aimed at supporting the needs derived from the growing use of digital tools and materials by social sciences and humanities scholars are falling short (Strange et al., 2023). Accordingly, the aim of this study is to investigate research data management practices of scholars conducting data-intensive social sciences and humanities research and propose a data lifecycle model based on their perceived needs from which improved research data services for these disciplines can be developed.

Research setting

Semi-structured interviews (n = 21) with researchers from different disciplines and various levels of experience in digital humanities and computational social sciences were conducted. Data collected were analysed through a qualitative content analysis focused on research data management practices. The interviews were carried out as part of the creation of a national research infrastructure for data-intensive social sciences and humanities research in Finland.

Findings

Results indicate that research data management practices in digital humanities and computational social sciences are impacted by unmet needs regarding both infrastructure and services, thus complicating the performance of these activities. Furthermore, scholars that conduct data-intensive social sciences and humanities research mostly carry out research data management practices at the beginning of the data lifecycle (collect, process, analyse).

Discussion

This study further supports the idea that researchers in digital humanities and computational social sciences not only engage to a certain extent in research data management practices, but existing research data services are also limited (Weller & Monroe-Gulick, 2014; Cox et al., 2019). Based on the perceived needs identified in the analysis, a community-based data lifecycle model (Carlson, 2014) specific to these disciplines is presented, including the main support needs required for each stage of the cycle. The model is conceptualized as an upward spiral following the recommendation of Cox and Tam (2018, p. 142), who suggested "to develop other metaphors and visualisations around research" beyond the lifecycle. We argue that the upward spiral illustrates in a better way the idea that research data management is a process that goes back and forth between the required tasks.

Conclusion

Enhancing research data management practices in digital humanities and computational social sciences depends both on the improvement of infrastructure and services related to these activities. In this context, the community-based data lifecycle model developed from the findings of this study provides a starting point for enhancing research data services aimed at supporting scholars conducting data-intensive social sciences and humanities research.

Acknowledgements

References

Borgman, C. L. (2007). *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. MIT Press.

Carlson, J. (2014). The Use of Life Cycle Models in Developing and Supporting Data Services. In J. M. Ray (Ed.), *Research data management: Practical strategies for information professionals* (pp. 63–86). Purdue University Press.

Cox, A. M., & Tam, W. W. T. (2018). A critical analysis of lifecycle models of the research process and research data management. *Aslib Journal of Information Management, 70*(2), 142–157. https://doi.org/10.1108/AJIM-11-2017-0251

Cox, A. M., Kennan, M. A., Lyon, L., Pinfield, S., & Sbaffi, L. (2019). Maturing research data services and the transformation of academic libraries. *Journal of Documentation, 75*(6), 1432–1462. https://doi.org/10.1108/JD-12-2018-0211

Khan, N., Thelwall, M., & Kousha, K. (2023). Data sharing and reuse practices: disciplinary differences and improvements needed. *Online Information Review, 47*(6), 1036–1064. https://doi.org/10.1108/OIR-08-2021-0423

Nahotko, M., Zych, M., Januszko-Szakiel, A., & Jaskowska, M. (2023). Big data-driven investigation into the maturity of library research data services (RDS). *The Journal of Academic Librarianship, 49*(1), 102646. https://doi.org/10.1016/j.acalib.2022.102646

Poole, A. H. (2016). The conceptual landscape of digital curation. *Journal of Documentation, 72*(5), 961–986. https://doi.org/10.1108/JD-10-2015-0123

Strange, D., Gooch, M., & Collinson, A. (2023). Equality, findability, sustainability: the challenges and rewards of open digital humanities data. *International Journal of Performance Arts and Digital Media, 19*(3), 348–368. https://doi.org/10.1080/14794713.2023.2206286

Weller, T., & Monroe-Gulick, A. (2014). Understanding methodological and disciplinary differences in the data practices of academic researchers. *Library Hi Tech, 32*(3), 467–482. http://dx.doi.org/10.1108/LHT-02-2014-0021