

INFORMATION STUDIES DAYS 2024



The use of digital images as research data: learnings from the ImAccess project

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>

Keywords: data acquisition, image search, pictures, research data

Pysyvä osoite: <https://doi.org/10.23978/inf.148600>

Visual data, such as images and photographs sourced from social media platforms and archives, are important empirical data for social sciences and humanities (SSH) scholars (Chassanoff, 2018; Chen et al., 2021; Rose, 2022). However, earlier research on data use in SSH has concentrated mainly on textual materials, resulting in a gap in understanding of interactions with image data. Images convey information differently from text, and impediments, such as copyright issues, may limit the utilization of images for research purposes (Rose, 2022). The aim of the ImAccess project (started at Tampere University in fall 2022) has been to investigate the use of images as research data in various fields of SSH to better understand users' real-life image data practices and needs (Late & Kumpulainen, 2022).

The first part of the project focused on the use of Finnish War Time photograph archive containing around 160.000 historical photographs. For the purposes of the study, we interviewed 21 expert users of the digital archive to learn more about how images are searched from the collection, what barriers users face while searching and what desires they have for developing image collections. The results show that users apply and combine different tactics (keywords, filtering and browsing) for image searching and rarely using one tactic only is enough (Late et al., 2024a). During searching users face various barriers, most of them related with keyword searching due to the shortcomings of image metadata. Our findings suggest that users' desires for the collection relate to three contexts: tools, collection, and socio-organizational issues indicating that users require support for various information interaction activities, not just searching (Late et al., 2023a). We found that users' desires vary based on their specific use purposes, and that users prioritize conceptual access points that can already mostly be generated through automated annotation methods.

As part of the project, a prototype tool was developed to search digitized historical images based on their visual content (Late et al., 2023b). Fifteen participants tested the tool and assessed their user experience using the User Engagement Scale (O'Brien, 2018), along with providing verbal feedback. The results show that the users benefited from the tool's search capabilities, which went beyond relying on textual image descriptions. However, challenges were identified, particularly in evaluating search results and user proficiency. The findings also highlight the importance of intellectually produced metadata for image searching and use. Therefore, future improvements should focus on developing hybrid systems that support both textual and visual image searching.

The second part of the project, that is still on progress, focuses on the use of more contemporary image data such as images and photographs sourced from social media platforms and archives. For the purposes of the study, we interviewed 21 SSH scholars using digital and print images collected from external sources as their empirical research data. Three

topics will be investigated: image data interactions, image data sharing, and collaborative activities. The findings will contribute to the task-based information interaction research for meeting the realities of interacting with image data.

Our analysis on image data sharing shows that while data sharing is encouraged as a part of the open science movement, it is not an established research practice, and when it happens, it is mostly done via informal means by sharing data through personal contacts (Late et al. 2024b). Data sharing is impeded by factors relating to the qualities of data, ownership of data, data stewardship, and research integrity. Therefore, advancing image data sharing requires the development of research infrastructures and providing support and guidelines.

Our future study will investigate the collaborative activities in image data interaction based on the model provided by Shah (2014). As previous studies on research collaboration have mostly focused on collaborations between team members, our aim is to provide more holistic understanding about the issue. The preliminary analysis shows the various types of collaborative activities taking place through-out the data interaction process from planning to concluding. Our data also indicate the importance of different actors, not just team members, in image data interaction. These include actors external to the project, such as other colleagues, study participants, data providers, publishers and research support services. Their role varies depending on the type of collaborative activity.

During the two years in the ImAccess project, we have learned that SSH scholars have benefited from the increased availability of image data and efforts put into the digitation of cultural heritage collections. However, providing digitized contents openly online is not enough if there are no sufficient means for accessing and using them. Automatic annotation methods are one option for creating metadata to improve the findability of the images. However, for enhancing the use of image data, scholars need support and collaborations throughout the whole image data interaction process. This includes legal and ethical information, method development and training and availability of research infrastructure.

Acknowledgements

This work was supported by the Research Council of Finland, grant number 351247. We thank Hille Ruotsalainen, Mette Skov, Inés Matres and Anna Sendra for their valuable contributions to the study.

References

- Chassanoff, A. M. (2018). Historians' experiences using digitized archival photographs as evidence. *The American Archivist*, 81(1), 135–164. <https://doi.org/10.17723/0360-9081-81.1.135>
- Chen, Y., Sherren, K., Smit, M., & Lee, K. Y. (2021). Using social media images as data in social science research. *New Media and Society*, 25(4), 849–871. <https://doi.org/10.1177/14614448211038761>
- Late, E., & Kumpulainen, S. (2022). Open image collections – sources of research data? *Informaatiotutkimus*, 41(2–3), 88–91. <https://doi.org/10.23978/inf.122591>
- Late, E., Ruotsalainen, H., & Kumpulainen, S. (2023a). In a perfect world: exploring the desires and realities for digitized historical image archives. *Proceedings of the Association for Information Science and Technology*, 60(1), 244–254. <https://doi.org/10.1002/pr2.785>
- Late, E., Ruotsalainen, H., & Kumpulainen, S. (2024a). Image searching in an open photograph archive: search tactics and faced barriers in historical research. *International Journal on Digital Libraries*. <https://doi.org/10.1007/s00799-023-00390-1>
- Late, E., Ruotsalainen, H., Seker, M., Raitoharju, J., Männistö, A., & Kumpulainen, S. (2023b). From Textual to Visual Image Searching: User Experience of Advanced Image Search Tool. In O. Alonso, H. Cousijn, G. Silvello, M. Marrero, C. Teixeira Lopes, S. Marchesin (Eds.), *Linking Theory and Practice of Digital Libraries*. TPD L 2023. Lecture Notes in Computer Science, vol 14241. Springer, Cham. https://doi.org/10.1007/978-3-031-43849-3_24
- Late, E., Skov, M., & Kumpulainen, S. (2024b). To share or not to share? Image data sharing in the social sciences and humanities. *Information Research an International Electronic Journal*, 29(2), 386–400. <https://doi.org/10.47989/ir292834>
- O'Brien, H. L., Cairns, P., & Hall, M. (2018). A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form. *International Journal of Human-Computer Studies*, 112, 28–39. <https://doi.org/10.1016/j.ijhcs.2018.01.004>
- Shah, C. (2014). Collaborative information seeking. *Journal of the Association for Information Science and Technology*, 65(2), 215–236. <https://doi.org/10.1002/asi.22977>
- Rose, G. (2022). *Visual Methodologies: An Introduction to Researching with Visual Materials* (Fifth Edition). SAGE Publications Ltd.