

NECS Pre-Conference Workshop:

Taking Care of Audiovisual Research Data in Film and Media Studies, Audiovisual Collections and Libraries

Scholars in the field of film and media studies have traditionally engaged in a broad range of practices that produce and involve different types of research data and personal databases. For decades, the annotation and segmentation of media items for formal and content analysis, the creation of filmographies as well as databases pertaining to distribution and reception of media have been integral to research in our fields. In recent years, such practices have gained prominence with the digitization of legacy data that form the basis for digital scholarship (Long, 2016; Drucker, 2021), or born-digital collections that are analyzed with digital methods at different scales, allowing for alternating between ‘close’ and ‘distant’ reading or viewing (Acland and Hoyt 2016; Arnold and Tilton 2019).

This development raises several questions and challenges as to how to take care of research data, both during and after a research project. For instance, the creation of such datasets involves the enrichment of audiovisual and textual items with various types of data, that could potentially also be used by other researchers (Melgar Estrada et al. 2017; van Noord et al. 2021), raising the question of how to publish and share research data with peers, which technologies to rely on, and the tricky problem of how to host and maintain ‘living’ datasets that are constantly updated and enriched. Likewise, the analysis of datasets consisting of data from legacy as well as born-digital sources requires a source critical mindset, and for allowing others to check and reuse data mindful of changing approaches and discursive underpinnings of data creation over time. For the institutions hosting such data – audiovisual collections and libraries - such questions are also becoming increasingly pressing, and in recent years especially the FAIR principles (Findable, Accessible, Interoperable, Reusable) have gained prominence as a reference frame for research and collection data management (Wilkinson et al. 2016).

Addressing this development, the NECS pre-conference workshop “Taking Care of Audiovisual Research Data”, co-organised with the National Library of Norway, presents, explores and critically evaluates different approaches. In the workshop’s first part, a number of short presentations will demonstrate examples of workflows used in recent film and media research projects, and of how to publish one’s data in a way that facilitates their discovery and

reuse by peers, considering platforms such as GitLab, Jupyter Notebooks, and Zenodo and Linked Open Data (LOD) approaches. These presentations are followed by a hands-on session in which participants can themselves explore the approaches discussed, and, where relevant, work with their own dataset. Finally, the workshop is concluded with a discussion and evaluation of the proposed workflows and hands-on exercises.

A detailed workshop program is available here:

<https://digital-methods-necs-workgroup.github.io/>

References

- Acland, Charles R, and Eric Hoyt, eds. *The Arclight Guidebook to Media History and the Digital Humanities*. Falmer: Reframe Books, 2016. <http://projectarclight.org/book/>.
- Arnold, Taylor, and Lauren Tilton. “Distant Viewing: Analyzing Large Visual Corpora.” *Digital Scholarship in the Humanities* 34, no. Supplement_1 (2019): i3–16. <https://doi.org/10.1093/llc/fqz013>.
- Drucker, Johanna. *The Digital Humanities Coursebook: An Introduction to Digital Methods for Research and Scholarship*. London: Routledge, 2021.
- Long, Derek. “Excavating Film History with Metadata Analysis: Building and Searching the ECHO Early Cinema Credits Database”. In *The Arclight Guidebook to Media History and the Digital Humanities*, edited by Charles R. Acland and Eric Hoyt. Falmer: Reframe Books, 2016. <http://projectarclight.org/book/>.
- Melgar Estrada, Liliana, Eva Hielscher, Marijn Koolen, Christian Gosvig Olesen, Julia Noordegraaf, and Jaap Blom. “Film Analysis as Annotation: Exploring Current Tools.” *Moving Image: The Journal of the Association of Moving Image Archivists* 17, no. 2 (2017): 40–70. <https://doi.org/10.5749/movingimage.17.2.0040>.
- Noord, Nanne Van, Christian Olesen, Roeland Ordeman, and Julia Noordegraaf. “Automatic Annotations and Enrichments for Audiovisual Archives.” In *Proceedings of the 13th International Conference on Agents and Artificial Intelligence - Volume 1: ARTIDIGH*, 633–40, 2021. <https://www.scitepress.org/PublicationsDetail.aspx?ID=DgZjO6HYYTA=&t=1>.

- Ordelman, Roeland, Willemien Sanders, Richard Zijdemann, Rana Klein, Julia Noordegraaf, Jasmijn van Gorp, Mari Wigham, and Menzo Windhouwer. "Data Stories in CLARIAH: Developing a Research Infrastructure for Storytelling with Heritage and Culture Data." *DARIAH Annual Event 2022*, May 31, 2022.
<https://doi.org/10.5281/zenodo.6597110>.
- Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. "The FAIR Guiding Principles for Scientific Data Management and Stewardship." *Scientific Data* 3, no. 1 (March 15, 2016): 160018. <https://doi.org/10.1038/sdata.2016.18>.