

# A Feminist Approach to Linked Open Data: Making the Women Film Pioneers Project FAIR

**Junginger, Pauline**

pauline.junginger@uni-marburg.de  
Philipps-Universität Marburg, Germany

## Introduction

The *Women Film Pioneers Project* (WFPP) is an established online resource for research on women in early cinema. The project focuses on telling individual stories of women film pioneers and making blind spots visible in order to draw attention to the major role played by women in the early days of film history (Gaines et al. n.d.). Initially conceived in the early 1990s as a classic book project, in 2013 the WFPP was published as a digital platform with three primary goals: to make research findings available, to stimulate further research, and to challenge traditional film historical narratives that fail to mention women's contributions. However, although structured metadata is central to making research data and digital publications visible, accessible, and reusable (Baca 2016; Flanders / Jannidis 2018), WFPP's application of metadata has been very limited. Therefore, my project aims to create structured metadata for the WFPP, enrich it with authority data, and link it in a sustainable way to other datasets on women film pioneers.

## Related work

In addition to feminist film historiography, the digital humanities are a central point of reference for this project, particularly approaches that address accessibility of text as data, digital publishing, categorization practices, and metadata (Blaney et al. 2021; Flanders / Jannidis 2018). In order to make metadata interoperable in the sense of the FAIR data principles, it can be made available as Linked Open Data (Schmidt et al. 2022). In the field of film heritage institutions, the modeling of film data as Linked Open Data has been discussed for several years (Heftberger 2019), including case studies with Wikidata (Heftberger / Duchesne 2020).

## Methodology

While the initial goal of the project was to develop a thesaurus tailored to the needs of WFPP's staff and users, it became clear as the project progressed that a more promising approach would be to focus on reusing existing thesauri to make the data FAIR. This conceptual shift from developing a project-specific thesaurus to opening up the collection by sustainably linking disparate datasets on women film pioneers had several implications for the project's methodological design. First, the nine categories initially identified for modeling the thesaurus (person, film work, occupation,

place, technique, institution, film genre, theme, and publication) were reduced to four categories relevant to the structured collection of filmographies: person, film work, institution, and place. By creating FAIR metadata using the WFPP filmographies as an example, the project explores how the FAIR principles can be applied in feminist film historiography. Second, whereas the initial concept included a manual text analysis and annotation with CATMA to build the thesaurus in a bottom-up approach, the filmographic data is now extracted using Named Entity Recognition and then modeled as LOD. Third, an important step in the methodology is to analyze and evaluate existing thesauri and authority files such as the Getty Art & Architecture Thesaurus (AAT), Friend of a Friend (FOAF), Virtual International Authority File (VIAF), Gemeinsame Normdatei (GND), and Wikidata regarding their usability for feminist film historiography.

## Critical reflection

From a feminist perspective, when working with Linked Open Data, it is important to ask what assumptions, biases, and power relations are inscribed in Semantic Web technologies as well as in authority files and vocabularies. Furthermore, it also involves an examination of the epistemic conditions of a structured data approach and a reflection on the specific requirements of feminist film historiography for applying the FAIR data principles. How can the plurality of feminist film historiography be accounted for when modeling research data as LOD requires categorization and standardization, and thus a reduction of complexity? How can we incorporate feminist values into the design of categories if we assume that classification practices are never neutral but always entangled with political attitudes and values (Bowker / Star 1999; Drabinski 2013; D'Ignazio / Klein 2020)? In the context of feminist film historiography, we also have to ask how the numerous blind spots and knowledge gaps that exist due to lost sources and the limited documentation of "feminized labor" (Hill 2016: 5) can be represented in data-based approaches (Dang 2020).

This poster outlines the methodological approach of the project and presents preliminary results.

## Acknowledgements

This project is carried out within the framework of the BMBF-funded research group "Aesthetics of Access. Visualizing Research Data on Women in Film History (DAVIF)".<sup>1</sup>

## Notes

1. Project website: <https://www.uni-marburg.de/en/fb09/institutes/media-studies/research/research-projects/davif> [01.05.2023].

## Bibliography

**Baca, Murtha** (2016): "Intro to Metadata", in: Baca, Murtha (ed.): *Introduction to Metadata*. 3rd ed. Los Angeles: Getty Publications. <http://www.getty.edu/publications/intrometadata> [07.07.2022].

**Blaney, Jonathan / Winters, Jane / Milligan, Sarah / Steer, Martin** (2021): *Doing digital history*. Manchester: Manchester University Press.

**Bowker, Geoffrey C. / Star, Susan Leigh** (1999): *Sorting Things Out: Classification and its Consequences*. Cambridge: MIT Press.

**Dang, Sarah-Mai** (2020): “Unknowable Facts and Digital Databases: Reflections on the Women Film Pioneers Project and Women in Film History”, in: *Digital Humanities Quarterly* 14, 4 <http://www.digitalhumanities.org/dhq/vol/14/4/000528/000528.html> [04.11.2022]

**D’Ignazio, Catherine / Klein, Lauren F.** (2020): *Data Feminism*. Cambridge: MIT Press.

**Drabinski, Emily** (2013): “Queering the Catalog: Queer Theory and the Politics of Correction”, in: *The Library Quarterly* 83, 2: 94–111. DOI: 10.1086/669547.

**Flanders, Julia / Jannidis, Fotis** (eds.) (2018): *The Shape of Data in the Digital Humanities: Modeling Texts and Text-based Resources*. London: Routledge. DOI: 10.4324/9781315552941.

**Gaines, Jane / Vatsal, Radha / Dall’Asta, Monica** (eds): *Women Film Pioneers Project*. New York: Columbia University Libraries <https://wfpp.columbia.edu/> [04.11.2022].

**Heftberger, Adelheid** (2019): “Building Resources Together – Linked Open Data for Film Archives”, in: *Journal of Film Preservation* 101: 65-73 <https://www.proquest.com/docview/2317838940/abstract/B700AB2C594C4879PQ/1> [01.05.2023]

**Heftberger, Adelheid / Duchesne, Paul** (2020): “Cataloguing Practices in the Age of Linked Open Data: Wikidata and Wikibase for Film Archives”, in: International Federation of Film Archives (ed.): *E-Resources* <https://www.fiafnet.org/pages/E-Resources/Cataloguing-Practices-Linked-Open-Data.html> [01.05.2023]

**Hill, Erin** (2016): *Never Done: A History of Women’s Work in Media Production*. New Brunswick: Rutgers University Press.

**Schmidt, Sophie C. / Thiery, Florian / Trognitz, Martina** (2022): “Practices of Linked Open Data in Archaeology and Their Realisation in Wikidata”, in: *Digital* 2, 3: 333–64. DOI: 10.3390/digital2030019.