

ARTICLE

Digital Film Historiography: Challenges of/and Interdisciplinarity

Malte Hagener¹, Diana Roig-Sanz^{2a}

¹ Institute for Media Studies and Marburg Center for Digital Culture and Infrastructure at the Philipps University of Marburg, ² ICREA and Universitat Oberta de Catalunya

Keywords: digital film studies, film historiography, digital humanities, data-driven analysis

<https://doi.org/10.22148/001c.120944>

Journal of Cultural Analytics

Vol. 9, Issue 4, 2024

Digital methods were slow to make inroads into film studies and particularly into film historiography. There are a number of reasons for this situation of extended latency: as a multimodal object, film requires a complex methodology for analysis which, in its translation into computational and data-driven approaches, calls for an adequately multi-layered theoretical and methodological framework. As a legal entity, film is heavily protected by copyright which means that access is either not possible in lawful terms or requires lengthy negotiations. As a digital object, film comes in the form of large files that are not easy to handle and need considerable computing power. Finally, the sources relevant to the cinema's past were rather peripheral and the necessary infrastructures for them were non-existent for the longest time. Many of the sources have not been digitized and they are still not available online. This is especially true for data related to historically underrepresented agents in film history such as women and for data pertaining to film cultures beyond the Global North. All these factors contributed to a situation in which applying data-driven methods to cinema history as a method of investigation became a difficult challenge.

Digital methods also require an investment of considerable resources: personal, technical, and institutional. Using digital methods in order to understand cultural phenomena, as it happens in other related disciplines, requires training, funding, personnel, a large amount of time for obtaining reliable results, and serious considerations in order to achieve long-term durability of both results and data. Undoubtedly, the agenda of many disciplines in the humanities were somewhat reluctant to these challenges. And there is still no consensus regarding how digital humanities as a field may fit with the particularities of specific communities in the humanities (cinema history, but also digital translation history, digital anthropology, or archaeology). Thus, we formulate a claim for a broader understanding which allows us to see our similarities, but also our differences and how different we

^a Diana Roig-Sanz acknowledges the support of the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme that allowed to organize the conference "Rethinking Film History Through Global and Digital Approaches (Early 1020s – Early 1970s)" in the framework of the ERC project "Social Network of the Past. Mapping Hispanic and Lusophone Modernity, 1898-1959" (Grant agreement No. 803860). Most contributions here included were presented at this conference.

are when applying data-driven approaches. Indeed, what distinguishes most digital approaches in the humanities and, more specifically, in digital film historiography is that we rely on data that is not fully structured and/or it does not fully exist yet and we need to produce it in the first place.¹ Also, there is still a lack of standards that allow us to work in a consistent and coherent way as a community. Quantitative and data-driven approaches were not at the core of film studies which has a tradition rooted in qualitative analysis and film aesthetics.

Therefore, up until roughly ten years ago, forays into digital film historiography were unsystematic, intermittent, and rather haphazard. As humanities researchers, we still tend to think “small” in terms of what we examine, but we have now the opportunity to think “big”. Yet, in order to do that, we urgently need some theoretical and methodological reflections, as well as empirical case studies that allow us to set the grounds of digital historiographies and offer general standards for our community and neighboring ones. Thus, beyond the question of data availability, there are also other fundamental issues such as that of (self-)perception. A large part of our community still does not have the sense of working with data in the process of research and we need to raise awareness about the existence of data - what data is and can be and if these data can be systematized and processed. In this respect, this special issue argues that data exist in film historiography and that these data can be examined computationally in a productive manner. We also aim at discussing theoretical and methodological challenges when applying digital methods to film historiography. On the one hand, our special issue wants to shed light on the question of the value of the digital in film historiography and the challenges we encounter when applying digital humanities methods to film studies and film history more specifically. On the other hand, we aim at discussing how to apply methods, protocols, and standards that, up until some years ago, were more closely related to the natural sciences than to the humanities, and this has certainly become a challenge both in terms of epistemologies as in terms of their application.

Today, more and more projects and institutions make sources and data available (which is a good thing), but often they are not sufficiently aware of their potential to work collaboratively and the standards and norms that have been established in the information science and library community. This takes us directly to the question of data quality, but also to the important issue of the lack of specific training in digital film historiography. The field of digital humanities is fragmented, consisting of many small communities which are not always well connected with larger infrastructural systems. To give but one example: which sources, identifiers and data standards should be used for a film as a singular work when cleaning and enriching data?

¹ This is the case, when we have to model data from existing sources which is a process that arguably lies between digitization and creation.

Should one refer to the Internet Movie-Database (IMDb) which started as a grassroots project in the 1990s, but which is now owned by Amazon and therefore shows limited transparency and responsibility regarding data quality? Or should one link one's data to national solutions such as the "Gemeinsame Normdatei - GND" (integrated authority file), a well-established format in Germany which holds a high liability there, but which might be practically unknown elsewhere? Or should one resort to an evolving and bottom-up initiative such as Wikidata which is open to participation, but which is also open to changes in governance and erratic data quality? As long as there are no established norms and authority files that data should adhere to (and this is the current situation in the field of film historiography), one has to make hard decisions again and again because sources are scattered and not always well-defined and structured. Between the data producers, those who make the data available via specific infrastructures and the users that need quality-assured data for working with digital tools there remains, unfortunately, too often a gap that needs to be closed. Of course, repositories with films as data or metadata exist in considerable numbers on the internet such as the following institutions, initiatives and databases: the "Moving Image Archive", the "International Federation of Film Archives", the "European Film Gateway", and "Colonial Film", showing images from daily life in the British colonies.²

The elephant in the room here is, of course, so-called "artificial intelligence" and how to make use of its widespread availability in the past two years. We are now facing several discussions on how the use of AI will contribute (or not) to the democratization of research and we still do not know to what extent AI will be beneficial for digital film historiography or what new challenges we will need to overcome. The use of deep learning models that have become available for ordinary users in the past two years (for text: Chat GPT, Bard; for still images: Dall-E, Stable Diffusion, Midjourney; for moving images: Sora which is not yet available to ordinary consumers) and the implementation of self-trained models into all kind of routines promises new opportunities, yet it also throws into relief that the way these tools arrive at their results is not transparent at all. If we use these methods (and both the essays by Diecke/Paiva and by Oiva *et al* make innovative use of those), we need to be wary of questions of replicability and transparency that used to be hallmarks of scientific working for good reasons. Also, AI can be helpful to get access to dark digital archives and to preserve content, but ethics and copyright issues also come to the forefront and need to be considered.

² Please see: <https://archive.org/details/movies>; <https://www.fiafnet.org/pages/Training/Metadata-Management-in-Film-Archives.html>; https://www.europeanfilmgateway.eu/about_efg/contributing_archives; <http://www.colonialfilm.org.uk/>

Having sketched the situation that made working with data-driven and digital methods in the field of film history difficult for the longest time, we acknowledge that the field is in the process of rapid change. With the massive digitisation of archives, magazines, and newspapers, important source material has been made available online and is engendering new research through large amounts of metadata and digital objects. This is most clearly visible through the most important showcase, the *Media History Digital Library*³ and other national and transnational infrastructural projects such as *media/rep/*,⁴ which has focused so far on the accessibility of texts free of copyright, but which is extending its portfolio into historical material. However, it is also fair to acknowledge that digitisation of archives is only a fraction of what digital film historiography should be. Indeed, as proposed by Roth, there are three potential ways of understanding digital humanities research: as “digitized humanities”, the creation, curation and use of digitized datasets in human sciences and, to a lesser extent, social sciences”, as “numerical humanities, by putting the emphasis on mathematical abstraction and the development of numerical and formal models”; and as “humanities of the digital, by focusing on the study of computer-mediated interactions and online communities”.

Within this general framework, attempts to institutionalize the field of digital film historiography have been promoted and discussing methods and epistemologies have become important goals for the activities of networks such as HOMER (History of Moviegoing, Exhibition and Reception),⁵ the Special Interest Group “Digital Humanities and Videographic Criticism” of the North-American SCMS⁶ (Society for Cinema and Media Studies) and the work group “Digital Methods” of NECS⁷ (Network for European Cinema and Media Studies). These groups connect the growing number of researchers interested in digital methods and their activities have supported (trans)national film historiography endeavors of a digital nature such as the “European Cinema Audiences Project”⁸ and the rich resource “Cinema Context”.⁹ There is also a growing number of projects that aim at advancing digital methods in general and also explicitly deal with questions of film historiography such as the “Media Ecology Project” at Dartmouth College.¹⁰ There is also an initiative such as the “Women Film Pioneers Project” at

³ <https://mediahistoryproject.org/>

⁴ <https://mediarep.org/home>

⁵ <https://homernetwork.org/>

⁶ https://www.cmstudies.org/page/groups_digital

⁷ <https://digital-methods-necs-workgroup.github.io/>

⁸ <https://www.europeancinemaaudiences.org/research/>

⁹ <https://www.cinemacontext.nl/>

¹⁰ <https://mediaecology.dartmouth.edu/wp/>

Columbia University¹¹ that deals with the female presence in film production during the silent period, and the two projects that initiated the current special issue of the *Journal of Cultural Analytics*: the ERC project “Social Networks of the Past. Mapping Hispanic and Lusophone Literary Modernity” at the Global Literary Studies Research Lab of the Open University of Catalonia,¹² particularly focused on data-driven approaches applied to film criticism, film clubs, and the history of women in cinema history, and the Digital Cinema-Hub network at the universities of Marburg, Mainz, and Frankfurt (funded by the Volkswagen foundation) which set itself as its aim the implementation of data-driven methods in film studies.¹³ A growing interest is also visible in recent conferences such as “Rethinking Film History Through Global and Digital Approaches” (Barcelona, October 2022)¹⁴ and “Doing Digital Film History” (Marburg, November 2022).¹⁵ Likewise, as tools for image analysis and object recognition are also becoming more readily available - the Distant Viewing Toolkit¹⁶ and the Clariah Media Suite¹⁷ might be the most visible examples to date - this might well be a turning point for rethinking the methods, questions and results of film historiography in the digital era. Also, there have been more efforts to make data openly available so that it can help encourage further research and collaboration, but also to help storage and preservation. These developments are also visible in terms of publications. Many of them defined what constitutes digital humanities as a field (Terras et al.) or as a set of methods and tools, but scholarship has also tried to reflect about the sense of history in the digital age (Liu), about the potential of computational methods for the study of culture (Manovich, *Cultural Analytics*) or about the use of big, little and no data in our networked world (Borgman). In film studies, about a decade ago, some publications dealt with the editing patterns in relation to digitality in the work of Soviet filmmaker Dziga Vertov (Manovich, *Visualising Vertov*; Heftberger), while New Cinema History pioneered the combination of exhibition data with geospatial information (Klenotic). In 2016, the *Arclight Guidebook to Media History and the Digital Humanities* provided a much needed overview of approaches and methods, underlining the development of the Media History Digital Library. In 2021, *Digital Humanities Quarterly* published a focus section on digital humanities and film studies, specifically on annotations and the analysis of moving images. And the long awaited foray by Arnold and Tilton into what they call *Distant Viewing* can surely be seen as a watershed

11 <https://wfpp.columbia.edu/>

12 <http://globals.research.uoc.edu/erc/>

13 <https://www.uni-marburg.de/en/fb09/institutes/media-studies/research/research-projects/dici-hub>

14 <https://blogs.uoc.edu/in3/rethinking-film-history-global-digital-and-gender-perspectives/>

15 <https://www.uni-marburg.de/de/fb09/medienwissenschaft/aktuelles/termine/international-conference-doing-digital-film-history>

16 <https://github.com/distant-viewing/dvt>

17 <https://mediasuite.clariah.nl/>

moment because it serves both as an introduction to the field, while also demonstrating the validity and productivity of diverse methods. There are more noteworthy publications, especially in the last five years and some are forthcoming (Dang et al.).

Within this more general context, the purpose of this special issue is to contribute to the discussion of digital film historiography and identify the main fields of advancement for the discipline. Building on these advances, the special issue aims at charting, on the one hand, what is at stake in current debates on digital methods and film history and what are the most challenging issues in the field. On the other hand, we are presenting cutting-edge research that demonstrates where crucial advances in the field are happening and how the development is being advanced with specific case studies. As stated above, we want to use this watershed moment by putting together a number of articles that address some of the most crucial current questions that emerge in relation to the digital transformation of film history. Among them, we would like to underline the following ones: 1) theoretical and methodological challenges when applying digital approaches and digital methods to cinema history; 2) the creation of protocols and standards owing to the lack of good infrastructures and sources prepared for the automatic extraction of data and/or digital analysis; 3) the discussion about power relations and underrepresented objects of research such as women, Black or Indigenous people. 4) the ownership and control of data. 5) unavoidable biases because of missing data or because of the fact that data has previously been created by human beings and, in many archives, we may find several mistakes and inconsistencies in the transition from analog to digital; 6) the need of raising awareness about the fact of providing data that can be used interdisciplinary and that can be reused and/or further developed by other colleagues from our community or related ones. 7) specific training so that younger generations of scholars do depend less on data scientists and data visualizers. 8) time to consider and interpret preliminary results. 9) actions for storage and preservation to fight against technological obsolescence. 10) acknowledgement of the value of open data and data sharing by both our funding institutions and our community, who are still a bit reluctant in terms of sharing their data and a research process that might not be concluded.

To reflect all these themes either directly or indirectly, we have gathered important case studies that demonstrate the potential of digital methods for the future development of film and cinema historiography all over the world. We want to address both the community of digital humanists, alerting them to developments in film studies, as well as those interested in film historiography who should be made aware of potentials and opportunities in using digital methods. The general and more specific threads these case studies will show are the following.

Methods and scales

When surveying the articles that have been gathered, we come across a broad range of (digital) methods that are being employed in the case studies. It might indeed be that the diversity of approaches and the complexity of workflow best illustrates the maturity that the field is reaching. In fact, most articles do not just use one specific method, but often combine different approaches. The studies collected in this special issue employ different ways of scalable readings, or, to be more precise: text mining, topic modeling, data processing and data visualization, GIS mapping, and network analysis for the paper by Torres-Cacoullos and Senatorova, social network analysis in the case of Clariana-Rodagut and Cardillo, but also for Hagener and Blaschke, sentiment analysis for the paper by Diecke and Paiva, machine vision and machine learning for Oiva *et al.*, to name but the most prominent approaches. Indeed, taken together the articles show that many of the methods used in the wider field of digital humanities and cultural analytics have been incorporated into film history in productive ways. It is also true to acknowledge the need for interdisciplinary work in most cases, as our contributors come from different fields ranging from media studies and cinema history to data science and physics. In this respect, a question that comes to the fore forcefully, if we look at the overall issue, is the relation between close and distant reading: how do we apply mixed methods and in what relation do we put the digital and the hermeneutical? How do we come from data to multi-sided, nuanced, and complex statements, but also: how do we find the data necessary to answer the question that we are having? Interestingly, many of the articles alternate between different scales, employing both a micro- und a macro-perspective. There is a need for complex negotiations and scalable readings. We do not see, for good reasons, the construction of one rigid data model that is then used in just one way. Instead, there is a regular alternation of different approaches, most visible perhaps in the complex workflow of Josephine Diecke and Isadora Paiva or in the multi-sided analysis of Oiva *et al.* In terms of methods, Diecke and Paiva have applied sentiment analysis to film reviews. As the authors claim, “this is common in the field of computer science, but it has been mostly absent in film studies”. The emergence of Large Language Models (LLMs; ChatGPT was released to the wider public in November 2022) “has shown that their capacity to handle nuanced language could overcome some of the shortcomings of lexicon-based sentiment analysis”. Oiva *et al.* apply multidimensional vector embeddings to study the internal time of the films, and the external time of the years from 1945 to 1992. In short, the reader will find examples for the application of machine learning methods for object and face recognition, as well as for shot scale and editing.

Topics, Sources and Tools

A similar broad range can be found when surveying the topics that are being treated in the articles of this special issue. They range from the temporal dimensions of film, as in the exploration of Soviet newsreels across several decades (Oiva et al), the role of male fandom in silent era Spanish cinema, as seen through contemporary fan journals (Torres-Cacoullos and Senatorova), film exile and the vicissitudes of material traces and archives, as it is manifested in a specific collection (Klages), the ebbs and flows of film production in the German cinema from 1919 to 1939 investigated across networks (Hagener and Blaschke), the role of film reviews in relation to classical German cinema (Diecke and Paiva), and the marginalisation of women in Ibero-American film clubs (Clariana-Rodagut and Cardillo). Some of them explore the relationships between individual and collective histories, and deal with historically marginalised subjects and topics such as film reviews, film clubs, the role of women in film culture, and geographies, as in the case of the Spanish-speaking world.

Similarly widespread are the sources used which shows that a broad range of material is publicly available, but also that material can still be unearthed in archives. Our case studies also show some of the challenges in the selection of the right and most relevant sources, the issue of data-collection and curation, as well as the efforts in cleaning and enriching the data with a great variety of sources, often scattered and heterogeneous, not always well-structured, or manually built. Hagener and Blaschke engage in the concept of what a national cinema is and what are their most appropriate sources. The discussion about sources also points to the multiple biases and inconsistencies we have to deal with when our datasets are not complete, not available or locked away. In this respect, the case studies are grounded in data and material found on both primary and secondary sources belonging to personal, local, national, and transnational physical and online archives: letters and correspondence, historical periodicals, trade journals and fan magazines in multiple languages, exhibition catalogs, photographs, online resources through the MHDL and archive.org, *Filmportal*, the central internet platform on German film, as well as data on people and films in the case of member lists or lists of screenings.

We cannot delve into all the challenges our contributors were confronted with regarding the quality of sources, but let us acknowledge a general issue that has to do with the fact of working with a big volume of material that was not digitized for large-scale analysis. In this respect, we would like to stress the important limitations when it comes to work with historical sources, unlike what occurs when researchers deal with born-digital material or when they can easily scrape it, as in the case of IMDb reviews. For example, Diecke and Paiva underline the difficulties of gathering the reviews for a large number of Weimar films through web scraping, so that a corpus can be created. In this respect, as Diecke and Paiva stress, this task is “far from

trivial" and they have created a corpus manually, despite being focused on a smaller sample of films. Also, film titles are "made up of common words, such as "Metropolis", "Passion" (the English title for Ernst Lubitsch's *Madame Dubarry*) or even letters, like "M", which make it difficult to identify the films without considerable manual work". Even more challenging has been "to distinguish reviews from ads (very positive) or plot summaries, [...] or to recognize when a review starts and ends in a text". Obviously, the quality of the Optical Character Recognition (OCR) is also a hurdle as it may both affect "the quality of the digital text but also the reliability of the search itself".

Workflows and infrastructure

Torres-Cacoullos and Senatorova ultimately aim at "a systematic survey of popular Spanish film magazines ". Yet again the many difficulties of processing the data results in a self-restriction to just one magazine for the article presented in this special issue. This is by no means meant as a criticism, but it rather shows that the employment of digital methods in the field of film historiography is still relatively early and much remains to be done, both in terms of making source material available as well as regarding the application of digital methods. The expectation of quick progress and low-hanging fruit that just need to be picked is often misleading. Imme Klages takes a different approach, as her starting point is less a concrete research question, but rather an archival collection that is turned into a data set. As data needs to be connected, the article traces the difficulties, but also the potentialities that linked open data offer for film history. Digital methods are, therefore, not only a tool to generate knowledge, but their epistemic potential can also be found in their power to uncover seemingly lost connections. As the article aims at unearthing the interrelations of film exiles, many of whom have been delegated to the margins, using digital methods can also work towards a redemption and a showcasing of the gaps left by history.

A similar path can be traced in the joint contribution by Ainamar Clariana-Rodagut and Alessio Cardillo who start off from a well-known observation - that women are being marginalised in the cultural domain just as in many other fields -, but turn it on their head by proposing a specific digital method for making this marginalisation visible. More specifically, they put forward "a methodological proposal combining quantitative tools of social network analysis with a feminist perspective in the field of historical cultural phenomena". By employing a novel way to ascertain centrality to a network (k core decomposition), they are putting forward a methodological proposal, but they also contextualise and critically reflect the productivity of this approach. Here, as in other essays showcased in this special issue, the cultural analysis done digitally needs careful framing, as well as specific domain knowledge in order to do justice to the complexities and multi-layered nature of cultural phenomena.

The approach chosen by Malte Hagener and Theresa Blaschke investigates a much larger data set - the national film production of Germany over a period of twenty one years (1919-1939). As is often the case, preparing the data is just as laborious as the analysis itself, but since the data analysis remains exploratory and is not functionally directed towards just one factor, the steps of finding and penetrating the data become an integral part of the analysis. Moreover, the preparation of rich and qualitatively good data also allows for asking other questions such as the centrality of specific crafts in the process of film production. Here, the tendency towards linked open data (and the general availability of data) becomes a driving force for expanding the range of questions and methods.

Another question that remains to be answered is the role of aesthetics for digital methods. Of the papers collected here, only Oiva et al use the films themselves as data. All the other essays are concerned with metadata about the films or with material surrounding the films, paratexts such as reviews or fan material. It will be interesting to see how the field will develop in the coming years, as more tools will be developed and become available that allow for easier and more options at working with the visual and acoustic material of the films themselves.

Concluding remarks

As stated above, the case studies included in this special issue clearly show the marked tendency for interdisciplinary collaboration which is nothing new to digital humanists, but which, so far, still remains the exception in film historiography. Only one of the articles that we include in this special issue is penned by a single author, all the others are the result of collaboration among different and multiple fields ranging from media studies and film history to computer science and physics. In this respect, the editors of this special issue are fully committed to an interdisciplinarity that it's not purely theoretical, but that we can take into practice, and that it's also well-acknowledged by funding institutions. Certainly, there is always the challenge of establishing collaborations with colleagues from other disciplines, working with a data scientist that is in our team, or promoting more training for junior staff. And we may also need to discuss if there already exists a sufficiently large number of methods and tools that can be appropriated and used in researching film history or if digital film historiography has its own specificities.

Without a doubt, film and media historians are becoming more familiar with data-driven approaches and quantitative analyses to examine and visualize with digital tools large corpora based on digitised archival sources that may reveal significant patterns and clusters that were not visible before. As pointed out at the beginning of this introduction, digital methods have the potential of making marginal groups visible, as more data becomes available in different parts of the world and through many initiatives. This is the case for data in relation to women's agency in film history, but also regarding minority

communities beyond the established centers and important nodes. Digital film historiography is on the rise and it seems only a matter of time until more research along these lines will be published and valued. Nevertheless, digital scholarship regarding cinema history still needs more established protocols, and a stronger institutional acknowledgement in the sense that many funding entities and our own community does not acknowledge sufficiently the value of the digital and the fact that some datasets are a scientific result in themselves. In this respect, we need to raise awareness about the value of this sort of research.

As editors of this special issue, we would like to stress that the editors of this special are fully committed to the important task of decolonizing cinema history and making more room available for those agents who were partly or completely excluded from mainstream and several national film historiographies. In this respect, we are committed to the FAIR and CARE principles, the latter ones not being represented in this special issue but implicitly acknowledged in our effort to promote a decolonial approach that can also shed light into less underrepresented agents such as Indigenous Peoples rights and interests.

This special issue has tried to expand the discussion about digital historiographies, primarily in cinema history, but also in other disciplines that work historically. We have also highlighted many of the challenges involved in digital film historiography (the wide range of sources, often scattered and non-structured, the investment of time in large-scale research, the issue of digital infrastructures, as in the case of Hagener and Blaschke, or the representation of forgotten actors in mainstream film history, as in the case of Clariana and Cardillo). We have also stressed the need to encourage multiple institutions, including film archives, libraries and public and private collections, to open their data, so that our historical past can be more accessible to society. Certainly, all these agencies are key actors that can encourage openness (open science, open access, open source, open peer review, but also open exchange, open borders), collaboration (mixed teams), and data sharing, but they can also contribute to produce data and to help granting more resources (financial, personnel, technical). Thus, we need to encourage film institutions to create and structure data in a way that the community can use them. That said, we will still need to push forward the reflection related to ethics and the ownership of data or the discussion on who creates knowledge in this specific field of research. But this will remain for a future discussion.

Submitted: July 14, 2024 EST. Published: July 18, 2024 EST.



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0/> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

WORKS CITED

Acland, Charles R., and Eric Hoyt, editors. *The Arclight Guidebook to Media History and the Digital Humanities*. Reframe, 2016, <https://projectarclight.org/book/>.

Arnold, Taylor, et al., editor. *Digital Humanities Quarterly*, vol. 15, no. 1, 2021.

Arnold, Taylor, and Lauren Tilton. *Distant Viewing. Computational Exploration of Digital Images*. MIT Press, 2023, <https://doi.org/10.7551/mitpress/14046.001.0001>.

Borgman, Christine I. *Big Data, Little Data, No Data: Scholarship in the Networked World*. MIT Press, 2015, <https://doi.org/10.7551/mitpress/9963.001.0001>.

Dang, Sarah-Mai, et al., editors. *Doing Digital Film History. Concepts, Tools, Practices*. De Gruyter, 2025.

Heftberger, Adelheid. *Digital Humanities and Film Studies. Visualising Dziga Vertov's Work*. Springer, 2018, <https://doi.org/10.1007/978-3-030-02864-0>.

Klenotic, Jeffrey. "Putting Cinema History on the Map – Using GIS to Explore the Spatiality of Cinema." *New Explorations in Cinema History. Approaches and Case Studies*, edited by Richard Maltby et al., Blackwell, 2011, pp. 58–84, <https://doi.org/10.1002/9781444396416.ch3>.

Liu, Alan. *Friending the Past. The Sense of History in the Digital Age*. University of Chicago Press, 2018, <https://doi.org/10.7208/chicago/9780226452005.001.0001>.

Loist, Skadi, and Evgenia (Zhenya) Samoilova. "How to Capture the Festival Network: Reflections on the Film Circulation Datasets." *NECSUS European Journal of Media Studies*, vol. 1, 2023, pp. 363–90, <https://doi.org/10.25969/mediarep/19615>.

Manovich, Lev. *Cultural Analytics*. MIT Press, 2020, <https://doi.org/10.7551/mitpress/11214.001.0001>.

---. *Visualising Vertov*. 2013, http://manovich.net/content/04-projects/078-visualizing-vertov/74_article_2013_sm.pdf.

Roth, Camille. "Digital, Digitized, and Numerical Humanities." *Digital Scholarship in the Humanities*, vol. 34, no. 3, Sept. 2019, pp. 616–32, <https://doi.org/10.1093/llc/fqy057>.

Terras, Melissa, et al. *Defining Digital Humanities. A Reader*. Ashgate, 2013, <https://doi.org/10.4324/9781315576251>.

Verhoeven, Deb, et al. "Controlling for Openness in the Male-Dominated Collaborative Networks of the Global Film Industry." *PLOS ONE*, vol. 15, no. 6, e0234460, June 2020, <https://doi.org/10.1371/journal.pone.0234460>.