Data-Driven Research for Film History

Exploring the Jean Desmet Collection

Olesen, C.G.; Masson, E.; van Gorp, J.; Fossati, G.; Noordegraaf, J.

Link to publication

Citation for published version (APA):

Olesen, C. G., Masson, E., van Gorp, J., Fossati, G., & Noordegraaf, J. (2016). Data-Driven Research for Film History: Exploring the Jean Desmet Collection. *The moving image*, *16*(1), 82-105. https://doi.org/10.5749/movingimage.16.1.0082

DATA-DRIVEN RESEARCH FOR FILM HISTORY

CHRISTIAN GOSVIG OLESEN,
EEF MASSON, JASMIJN VAN GORP,
GIOVANNA FOSSATI, AND
JULIA NOORDEGRAAF

Exploring the Jean Desmet

Collection

Cinema Context (848)

Collection Eye (1082)

Jonkman (863)

899

GENRE

comedy (526)

drama (143)

romance (101)

thriller (81)

✦ Show all options

COLOR

Black and White (732)

Tinting (375)

Toning (43)

Colourization (18)

✦ Show all options

PRODUCTION COMPANY

Pathé Frères (118)

Gaumont (94)

Nordisk (49)

Cines (44)

✦ Show all options

164

54

Ri

149

70

2

Nationaal Park
Oosterschelde

54

110

2

A10

Brugge

E34

E40

E17

In the past decades, film archives have invested vast amounts of energy and resources into digitizing significant parts of their holdings, including collections of early cinema.[1] This steadily growing digital archive presents a treasure trove for researchers attempting to understand the early years of the medium and its role in modern societies. It also provides new opportunities for audiences to engage with cinema history in attractive ways—opportunities we have only just begun to mine.

At the same time, the exploration and analysis of these digital archives require new means, as does the presentation of results to audiences. Digital tools are helpful here because they can provide comprehensive overviews of collections and reveal macrolevel patterns and structures within them. They are useful starting points for the formulation of

new questions and for the development of innovative research paths.[2] However, the use of these tools also requires reflection on their underlying assumptions and limitations. Digitization transforms physical objects into data, sometimes at the cost of the information available in the original sources (e.g., in the case of film, edge mark information or material features that do not easily get captured in the course of digitization).[3] Online collection overviews tend to emphasize factual information such as dates and locations, privileging the known at the expense of the unknown and not always doing justice to a collection's variety.[4] Cultural heritage metadata vary enormously, both in extent and quality, and this may skew the interpretation of research results.

In the recent project "Data-Driven Film History: A Demonstrator of EYE's Jean Desmet Collection," researchers explored the potential and pitfalls of digital methods for research into early cinema history.[5] It was a small-scale project lasting eight months part-time that involved academic researchers at the Universities of Amsterdam and Utrecht, along with staff from EYE Filmmuseum and two technical partners.[6] It centered around the recently digitized collection of early film distributor and cinema owner Jean Desmet (1875–1956), which is part of EYE's renowned corpus of silent films and related materials. The collection includes films, photographs, posters, and Desmet's business archive and was digitized as part of the Images for the Future project (a large-scale digitization effort funded by the Dutch government between 2007 and 2014).[7]

From an academic perspective, the objective of the "Data-Driven Film History" project was to try out, and in one case to develop, tools for studying the distribution, screening, and stylistic features of the films in the Desmet Collection. In particular, we sought to test the usefulness of such software in establishing relations between the distribution and screening of the films, and between their screening or programming and some of their aesthetic qualities, specifically their use of color. For EYE, the objective was to gain a better understanding of the quality of its metadata on the Desmet Collection and to experiment with new ways of providing access to information about the films. The software partners were interested in contributing to technologies that could be used for research and access purposes within the heritage sector more broadly.

As part of the project, we developed a mapping tool and experimented with software for the visualization of color patterns in films. The mapping tool was designed to help us visually represent contextual information and map it onto the titles of films in the collection. The visualization software in turn was chosen to help us discern textual patterns in the films themselves, specifically chromatic ones. Combining these tools, we hoped, would allow us to suggest some new pathways for the study of colors in silent cinema and enable researchers to capitalize on the opportunities afforded by recent

digitization efforts. Along the way, the project also presented some intriguing method-ological issues, which provoked much of the reflection we present here. The process of building a mapping tool, and, in particular, the ancillary work of combining metadata originating from different sources—three different data sets, produced at different times and with different objectives—invited us to consider, among other things, the limitations of our data and how these might affect the historical insights these kinds of tools might help us gain in the future. In assessing the media historical merits of these software pack-ages, we pay particular attention to what the combination with a mapping might reveal.

THE DESMET COLLECTION

The Jean Desmet Collection, preserved at EYE in Amsterdam, contains the archives left behind by the Dutch film distributor and cinema owner Jean Desmet. As a distributor, Desmet was most active in the early period of silent cinema and its transitional years.[8] The collection consists of approximately 950 films produced between 1907 and 1916, a business archive of more than 100,000 documents, some 1,050 posters, and approxi-mately 1,500 photos.[9] In 1957, some months after Desmet's death, most of the collection was acquired by what was then known as Stichting Nederlands Filmmuseum.[10] In the decades following its acquisition, the collection became of key importance to research on early silent cinema. It is distinguished for its concentration of films from the tran-sitional phase of silent cinema as well as for the richness of the associated business archive, which uniquely documents the early practice of film exhibition and distribution in the 1910s. Because of its historical value, it was inscribed on UNESCO's Memory of the World register in 2011.[11]

The film collection's significance is due in large measure to the fact that re-searchers have over the years unearthed many films that were previously considered lost. This enabled film scholars to expand their filmographies and their knowledge of genres, directors, and production companies. Moreover, the rediscovery of films has nurtured a more profound insight into the characteristics of early cinema, such as its common genre conventions or the widespread use of colors. Until the 1980s, for instance, the fact that the majority of silent films applied a color process—in the form of tinting, toning, stenciling, or a combination of those—was largely ignored by archives and scholars. Prior to this date, the predominant archival screening practice had been to present preserved films in black and white, with little regard for the colorful appearance of their source elements. When, in the late 1980s, the Filmmuseum allotted more resources for the registration, restoration, and dissemination of the Desmet Collection, it decided

to attempt to reproduce the colors of the original elements. It began to circulate color prints at festivals, thus actively opposing established practice. Seminal screenings in this regard were those at the Pordenone Silent Film Festivals of 1986 (with a color print of the Italian diva film *Fior di Male,* Cinès, 1915) and 1987 (featuring color restorations of Vitagraph films). Regarded in retrospect as significant events for both film archival practice and the study of early cinema, the screenings highlight the Desmet Collection's role in yielding a fresh outlook on national cinemas, production companies, and their textual conventions.[12] The collection also is associated with a rich business archive containing Desmet's meticulous notes on his transactions with cinema traders and exhibitors. These sources allow researchers to see exactly where prints were acquired or rented out. Therefore they yield key insights into early cinema's distribution practices and the dynamics of the film trade in the early and transitional years, with implications for the study of film programming.[13]

The Desmet Collection now has been almost entirely digitized and invites different kinds of access and new types of analytical intervention. So far, the digitized collection primarily has been used for exhibition. New scholarly insights, however, will be gained by applying digital methods to study the collection as raw data. In the context of "Data-Driven Film History," we considered how contemporary forms of data visualization might be instrumental in reassessing our perspective on issues such as film distribution, exhibition, and the uses of color in cinema's early years.

DATA-DRIVEN RESEARCH

Our primary motivation for developing a digital tool to examine the Desmet Collection was that, in digital methods for film historical research, visual forms of representation prevail over written ones. Diagrams, maps, video essays, and network visualizations have gained a stronger foothold in the humanities, as a means not only for presentation but also for analysis. The traditional written word seems to be yielding some space to the (moving) image as a means for scholarly knowledge production.[14] Digital historian David J. Staley has argued that visual formats can offer qualitatively different ways of exploring and analyzing historical sources and, depending on one's purpose, may even be better suited to an understanding of the intricate time–space relations that constitute historical situations.[15] Visual forms of representation, he argues, can produce "thick depictions" of historical sources, as they convey large amounts of information simultaneously. An added advantage of such depictions is that they do not have a clear beginning and end point and therefore (still according to Staley) allow for a good amount

Figure 1. Handcart advertisement for the film *Dødsspring til hest fra cirkuskuplen* (Eduard Schnedler-Sørensen, Denmark, 1912) screened at the Cinema Parisien in Rotterdam, owned by Jean Desmet. From the Jean Desmet Collection. Image reproduced courtesy of EYE Filmmuseum.

of "syntactic freedom" on the researcher's part. They enable a multidimensional exploration of the relations between historical sources and events without a fixed entry point.[16] In other words, they enable us to discern patterns and regularities, without forcing us necessarily to interpret them as causation.

Against the benefits of each tool, of course, stand its limitations. For instance, Deb Verhoeven, Kate Bowles, and Colin Arrowsmith, addressing specifically the use of Geographic Information System (GIS) software in historical audience research, argue that while they may indeed visualize certain spatial patterns in cinema culture and cinemagoing, they tend to neglect the temporal features of historical events (and, specifically, how they unfold or develop over time).[17] However, what we were keen to discover in the context of the "Data-Driven Film History" project was how the use of such software could help *complement* scholarship based on traditional methods, by making visible, and thus drawing attention to, other (kinds of) patterns and relations. More specifically, we were keen to try out a combination of two types of visualization tools—a combination that would allow us to source both the digitized collection and the associated metadata EYE has collected or produced over the years.

On one hand, we set out to develop our own tool for the visualization of meta-data, scraped from different data sets on the distribution and exhibition of films in the collection, onto a map. Our main reason for choosing a mapping tool was that this format has proven highly productive in recent years for the analysis of cinema's socioeconomic networks of production, distribution, and reception. Film historians working in a tradition of New Cinema History, a tradition that focuses on precisely those aspects of the cinema industry and moviegoing, have expressed great excitement about the potential of maps to convey spatial relations between vast amounts of historical data.[18] In retrospect, we can attribute our choice to develop a Mapping Desmet tool (discussed in more detail later) to a similar desire to mine the richness of the geographical information in the data sets we had at hand.[19]

On the other hand, we also wanted to try out some computational tools for content analysis, specifically the kind that had previously proven productive in the detection of patterns in film color. We were convinced that deploying these might generate more insight also into color use in films from the Desmet Collection, one of the features that it is famous for. We were hoping in particular that it might fuel a reflection on how the geographical information used for the mapping might be productively combined with such forms of analysis.

As a team of humanities researchers and archivists, we thought it key to consider at an early stage in the project also the methodological implications of what we were about to do. Above all, we wanted to stay aware at all times that the visualizations our tools produced were not going to be transparent, direct representations of a reality and that the results they generated would be subject at least to some extent to interpretations already embedded in the data.[20] One way to accomplish this was to allude to this in our interface design.[21] To be able to openly acknowledge the ways in which the data and their visualization would inevitably affect any interpretation of our results, we decided to build this interface ourselves. In doing so, we followed Stephen Ramsay's motto that building tools as such can provide new insights: "Building is . . . a new kind of hermeneutic—one that is quite a bit more radical than taking the traditional methods of humanistic inquiry and applying them to digital objects."[22]

MAPPING DESMET: GEOLOCATING THE DISTRIBUTION AND SCREENING OF EARLY FILMS

Our main goal in developing the mapping tool was to visualize the geographical patterns behind the distribution and screening of films owned by Desmet so as to allow users to gain insight into whether, when, and where they were rented out and shown. By
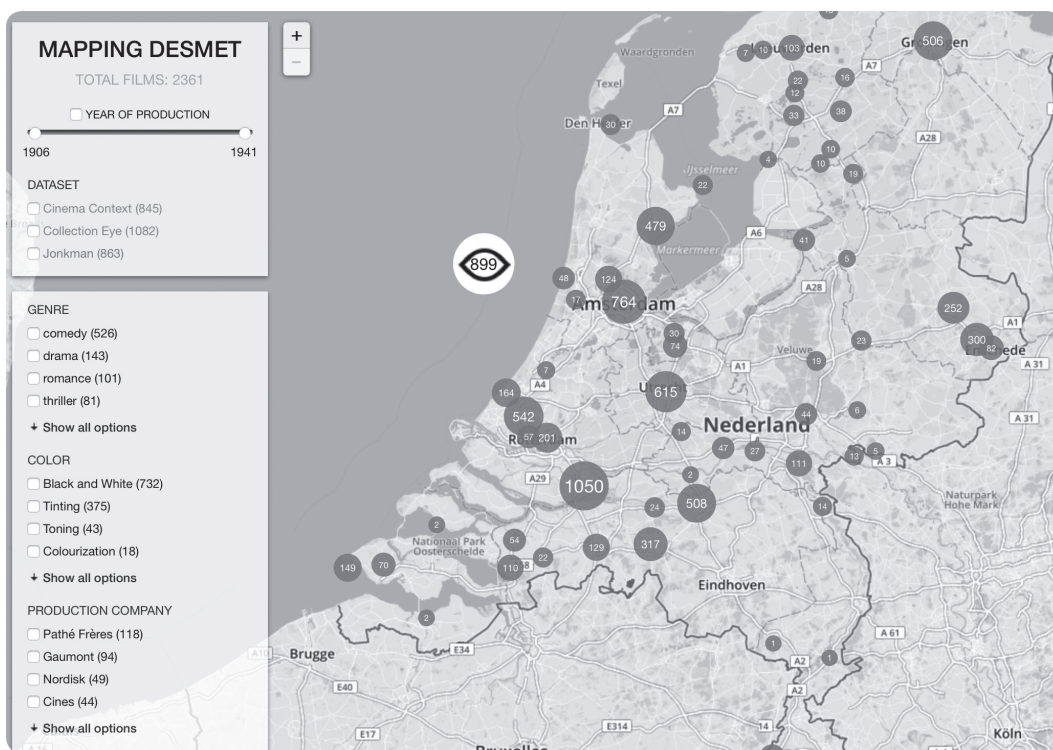
Figure 2. Screenshot of the default view of the demonstrator. The names of the data sets are color coded in the online version (http://mappingdesmet.humanities.uva.nl/#/map). Images reproduced courtesy of the University of Amsterdam.

providing macrovisualizations of their distribution and screening, vis-à-vis microdescriptions of the individual films, we aimed at reuniting all available information with regard to the collection in a single mapping that could facilitate the navigation of such an abundance of information but also inspire new research directions. One feature of the Desmet Collection in particular was key to our interface design: the fact that it is documented in three different databases, which vary in terms of the kinds and amounts of information they hold. It is the heterogeneity of the data sets deriving from those three databases that formed the starting point for the development of, and that became the backbone for, our mapping.

The first data set we used is derived from EYE's institutional catalog, which contains data for which the institution holds material (e.g., a film print, still, poster, or advertisement).[23] This is the most complete data set with regard to the Desmet Collection, as it comprises approximately 950 titles out of the 1,500 to 2,000 films Desmet acquired and distributed during his active years. Unfortunately, Collection EYE contains limited information on its distribution or screening.

The second data set is one produced by EYE collaborator Rixt Jonkman,

who made a detailed transcription of distribution and rental information from Desmet's business archive. Jonkman manually entered data on the rental, exhibition, and distribution of the 771 films mentioned in the account books that were purchased from two German distribution companies, Westdeutsche Film-Börse (WFB) and Deutsche Film Gesellschaft, between 1910 and 1912. These films were typically shown in "programs" containing a number of short films followed by a longer one (a circumstance we shall return to when discussing the color visualizations created as part of the project). While it is the most reliable source with regard to exhibition and distribution information, it only covers the period from 1910 to 1912, when Desmet bought his films primarily from Germany.

The third and last available data set is derived from Cinema Context, an online database developed by Dutch film historian Karel Dibbets. It relies on newspaper clippings and reports from the Centrale Commissie voor de Filmkeuring (Central Commission for Movie Ratings, installed in 1928) to establish where and when films from the collection were shown in the period covered by the collection. Cinema Context, however,

only includes information for one screening a week—typically the premiere screening—for each film title. It should be noted here that in 2004, data from the Jonkman database were added to Cinema Context, but in the meantime, Cinema Context has also been updated (e.g., unidentified film titles have been identified or Dutch distribution titles have been replaced by original titles). This is probably the main reason why not all the titles in the Jonkman data set match those in Cinema Context.[24]

To render visualizations from three different data sets, a strategy had to be found for coping with the many differences in terms of how they described the films in the collection. The titles and screening dates especially diverged considerably between the data sets, and therefore the items this information concerned could not be easily matched. For instance, the three data sets used titles in different languages (the original title versus the Dutch distribution title, or, in the case of an unidentified film, even a catalog title). By analogy, they also used different dates, as some considered the year of production, whereas others adopted the year of release or of the first screening—if they had dates at all. Although metadata specialists and curators at EYE made an effort to match a good deal of titles manually (in the process also "cleaning up" some of the records in Collection EYE), discrepancies in information and duplication in titles continue to exist.[25]

To draw the user's attention to the remaining discrepancies, we decided to follow Johanna Drucker, who, in her book *Graphesis,* advises her audience to visualize the contingencies and uncertainties associated with their data in the interfaces they use.[26] For us, this meant in practice that we tried to foreground the provenance of our data and, consequently, the limitations of the sources we used. Doing so should allow users of the demonstrator not only to assess the status of our data but also to determine what they can or cannot gain insight into when using them. While developing our tool, we implemented the principle of provenance on different levels.

Upon opening the demonstration tool, the user is shown a map of the Netherlands along with, in the upper left corner, a key naming the three data sets (see Figure 2). Checking the box next to one of them produces the total numbers of titles associated with that data set, mapped onto the cities in the Netherlands where they were distributed or shown. In the key, the data sets are color coded, and the same coding is used in the list of films that is shown when the user clicks the figure for a particular city (see Figure 3 for an example). Because we do not have details on distribution or screening for all of the films we know Desmet owned, we confirmed that those titles are also accessible through the interface.[27] Finally, at the title level, colored dots are used to show which data set specific information on genre, production, screening, and so on, derives from (Figure 4).
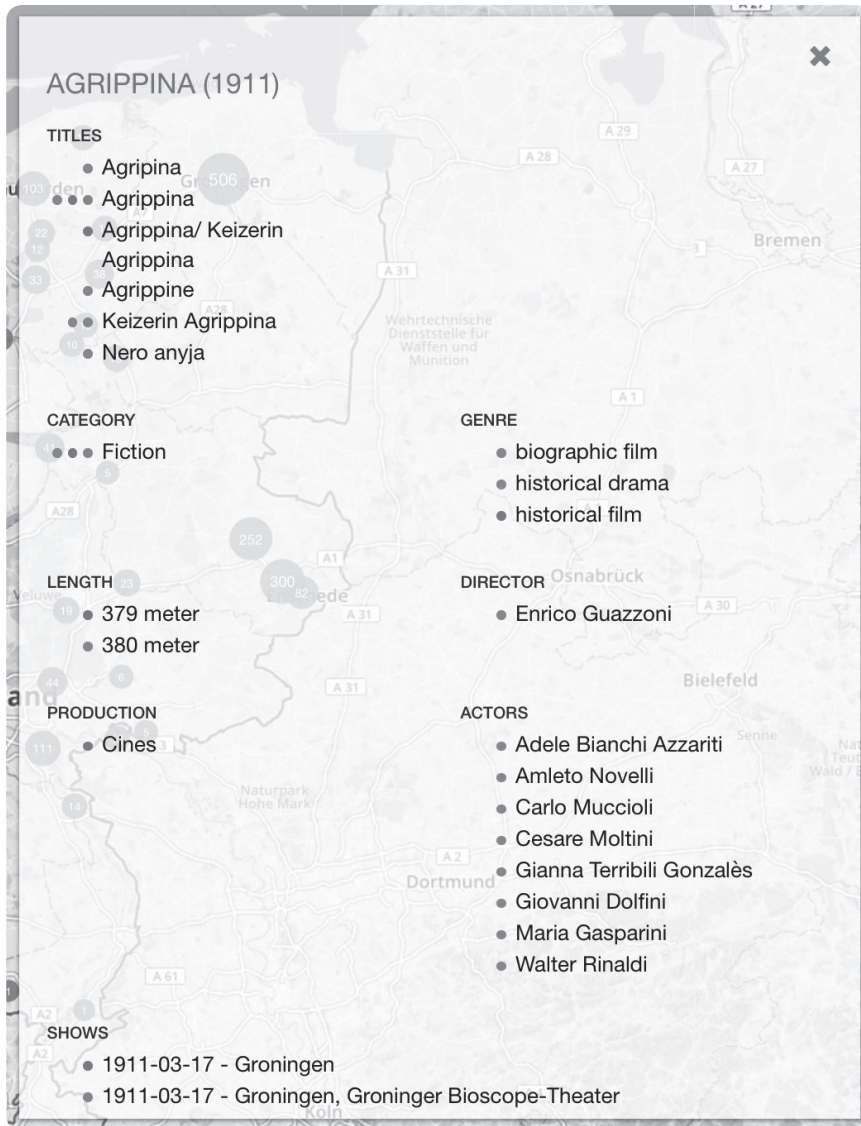
Figure 4. Screenshot of metadata fields and categories listed for the film *Agrippina* (1911). The dots, color coded in the online version of the demonstrator (http://mappingdesmet.human ities.uva.nl/#/mapitem /AVCEYTEq4tnq3xaYVOZB), identify the databases from which the information is derived. Image reproduced courtesy of the University of Amsterdam.

In the end, then, the availability of three data sets documenting the Desmet Collection turned out to be a decisive factor in the development of our mapping tool. By color coding each data set and making the provenance of the data visible in the demonstration, we sought to ensure that researchers stay aware of the fact that the information they use has been produced under different conditions and with different purposes in mind. The process encourages them to consider how this affects the relations they establish between the distribution and screening of the Desmet films, or any other film historical observations they might make. As discrepancies in data sets are issues archivists and archival researchers must deal with often, and in light of the fact that the resources to clean and merge such sets are rarely available, a visualization of the items' provenance can be a useful alternative, as it at the very least highlights problems that cannot be easily resolved.

ANALYZING COLORS IN THE DESMET PROGRAMS: TOWARD A CULTURAL ANALYTICS OF CHROMATIC PATTERNS IN EARLY CINEMA EXHIBITION?

In addition to developing a mapping tool, we were prompted by the Desmet Collection's significance for film color historiography also to experiment with tools for color analysis. This way, we hoped to create some of the preconditions for a more profound understanding of the palettes of silent cinema, specifically in the context of its exhibition. As pointed out by publications since the late 1990s, when the study of color in early cinema gained momentum, the majority of films in the silent period were probably shown in color (being hand painted, tinted, toned, or stenciled), and audiences expected colorful programs from exhibitors.[28] Taking this into consideration, we set out also to try our hand at color analysis, with the ultimate goal of linking our results to our mapping interface in some way. At the most basic level, we sought to experiment with ways of producing compact views of the sorts of color schemes audiences were exposed to when visiting any of the cinemas where Desmet's films were distributed and shown. In the future, we hoped, integration with our mapping tool should also allow users to relate the macroscopic perspective of a map and its visualization of distribution networks to the microscopic color features of a film or film program, through one and the same interface—a combination that is rare in digital scholarship to date.

In recent years, visualization software that has been widely used in scientific research practice is beginning to enter the repertoire of film studies analytical techniques—among those also software for color analysis. For example, film scholar and psychologist James Cutting has used the data analysis and visualization environment

Matlab to perform tentative analyses of color palettes in contemporary Hollywood films as a tool for studying audience perception.[29] In a technically similar fashion, but addressing a more aesthetic research interest, media theorist Lev Manovich's Software Studies Initiative has created ImagePlot, an extension of the scientific open source software ImageJ, initially conceived for medical imaging (among other purposes).[30] One of the advantages of ImagePlot is that its usability is not limited to the cinema. It can handle a wide array of visual media, such as paintings, graphic artwork, and photographs, and is specifically designed to work with large image sets. Therefore Manovich and others have designated the software as a "cultural analytics" tool; it suggests a methodological foundation for identifying changes in visual forms of cultural production over several decades and in different forms on a larger scale than hitherto possible. In the broader context of digital humanities approaches, cultural analytics can be characterized as a form of "distant reading," which allows for the visualization of the rise and fall of genres and the development of their key features over time, often as related to changes in production circumstances or concurrent historical events.[31]

In the context of our project, we chose to focus specifically on an analysis of the color palettes that Desmet's historical audiences would have been exposed to. To this end, we used digitized versions of three programs with films from Desmet's distribution catalog that had been screened in two cinemas in Rotterdam, in 1914, and one in the smaller, provincial town of Vlissingen, in 1915.[32] As Table 1 shows, each consists of a fairly conventional blend of genres for the time: a nonfiction film, a drama, a comedy, and so on. By making visual summaries of the hues featured in these programs, we hoped to get an overall impression of the presence, range, and/or distribution in time of the colors that contemporary audiences would have seen, an impression impossible to obtain with the naked eye. While the overviews thus produced are necessarily specific to a very selective number of historical places and times, and therefore cannot allow us to make any sort of generalizing statements, we hoped that, on the basis of a more extensive sample, they would at least permit us to consider the feasibility of making broader inferences on either the geographic distribution or the development over time of color use in silent film.

Again, methodological caution is in order here. We know that, historically, colors were "disorderly ordered." Nicola Mazzanti notes that even if we now tend to acknowledge colors and their patterns' role in silent cinema's processes of signification, we still know very little about the highly contingent factors that shaped them.[33] In retrospect, then, we cannot consider them fixed categories. Moreover, despite its extensiveness, the Desmet Collection hardly stands as a comprehensive representation

Table 1. Specifics of the film programs that served as the cases for our color visualizations.

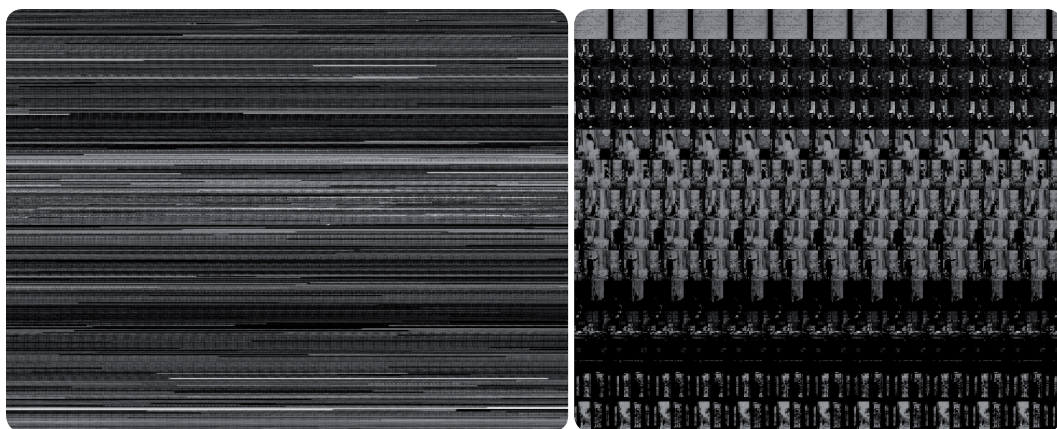|	*Program 1*	*Program 2*	*Program 3*
Program number, venue, date	De Gezelligheid, Goudseweg 124, Rotterdam	Concordia Bioscope, Schiedamscheweg 19, Rotterdam	Bellamy-Bioscoop, Bellamypark 18, Vlissingen
	20–26 February 1914	25–31 December 1914	1–7 January 1915
Films	Main program:	Program:	Program:
	1. *Amerikaansche vlootmanoeuvres* (*Target Practice of Atlantic Fleet* US Navy, Edison Manufacturing Company, USA 1912)	1. *Loetschberg* (*Le Chemin de fer du Loetschberg*, Eclipse, France 1913)	1. *Sous le ciel Basque* (*Onder de Baskische Hemel*, Eclipse, France 1913)
	2. *Verbroken geluk* (*Gebrochene Schwingen*, Adolf Gärtner, Messter, Germany 1913)	2. *Geduld overwint alles* (*When Persistency and Obstinacy Meet*, Vitagraph Company of America, USA 1912)	2. *Het eerste duel van Willy* (*Il Primo duello di Polidor*, Ferdinand Guillaume, Pasquali, Italy 1913)
	3. *Onwetend verloofd* (*Solitaires*, Van Dyke Brooke, Vitagraph Company of America, USA 1913)	3. *De avonturen van Fifi* (Eclipse, France 1913)	3. *Mijnheer Pijp champignon-kweeker* (*Mr. Pyp als Champignon-zuechter*, Charles Decroix, Monopolfilm, Germany 1913)
	Extras:	4. *Een autopech* (*Panne d'auto*, Baldassarre Negroni, Celio Film, Italy 1912)	4. *De slachtoffers van een woekeraard* (*Il focolare domestico*, Nino Oxilia, Savoia Film, Italy 1914)
	4. *Kwellende herinneringen aan het verleden* (*L'Obsession du souvenir*, Gaumont, France 1913)	5. *Overwinnen of sterven* (*Vittoria o morte*, Itala Film, Italy 1913)	5. *André op de planken* (*Andy Goes on the Stage*, Charles H. France, Edison Manufacturing Company, USA 1913)
	5. *Leon gaat naar buiten* (*Léonce à la campagne*, Léonce Perret, Gaumont, France 1913)	6. *De buurvrouw van Contran* (*Gontran et la voisine inconnue*, Eclair, France 1913)	6. *De orgeldraaier* (*The Organ Grinder*, Kalem, USA 1912)
	6. *De dochter van de Gouverneur* (*Guvernørens datter*, August Blom, Nordisk Films Kompagni, Denmark 1912)		7. *Polycarpe is aan het schijfschieten* (*Polycarpe veut faire un carton*, Ernest Servaes, Eclipse, France 1914)
Duration of digitized programs	01h51m37s	01h30m07s	01h46m28s

Figure 5. *Left,* black-and-white version of the ImagePlot montage visualization of Desmet program 1 and, *right,* a zoom image of the same. Full-color versions of these images can be accessed at http://www.create.humanities.uva.nl/#/. Images reproduced courtesy of EYE Filmmuseum/University of Amsterdam.

of the entirety of the production and circulation of films in the period. Last, but not least, anyone who has been involved in film archival practices will know how the duplication of film for restoration or access purposes, both analog and digital, can affect the rendering of the colors originally applied to black-and-white film.[34] For those reasons, we cannot, at this stage, consider color in these films in any other way beyond a very exploratory one, and also, very specifically, in relation to the Desmet Collection.

In our visualization work, we primarily used the ImagePlot software created by Manovich's Software Studies Initiative. On a practical level, this choice was motivated by the tool's thorough documentation, which makes it easy to use without much coding knowledge. In addition, ImagePlot is also one of the few visualization tools that was picked up by media scholars to study archival film and classic and contemporary cinema (albeit still on a small scale).[35] Having other scholarly projects as points of comparison allows us to situate our efforts within an ongoing scholarly debate on the potential of ImagePlot's techniques for developing new perspectives on film history.

For our purposes, each of the Desmet programs' video files was broken down into image sequences consisting of all the films' frames. In each case, this produced image sequences of more than one hundred thousand images.[36] From these sequences, we created different types of visualizations to find out which kinds of patterns each would reveal: first for the full programs and then for the individual films. We hoped this would allow us to reflect on the qualitative differences between them and their respective analytical potential.

A montage visualization of program 1, as shown in Figure 5, rearranges its frames into successive horizontal rows, to be read from left to right and top to bottom.
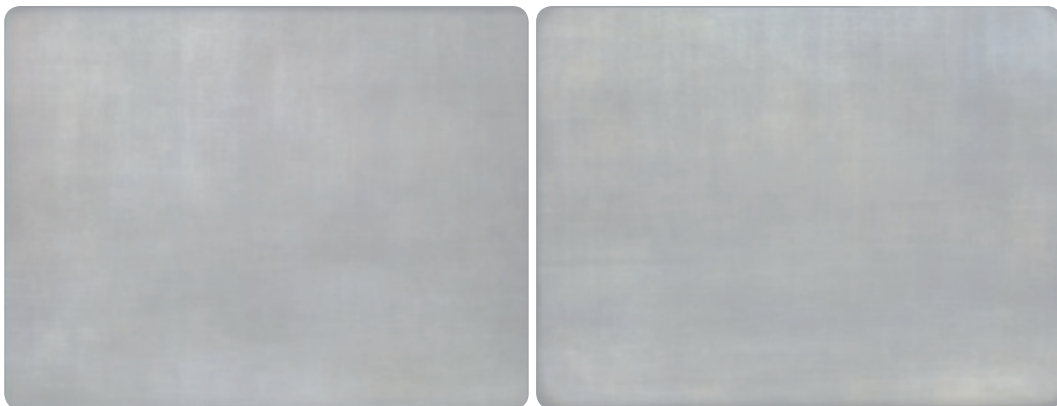
Figure 6. Black-and-white versions of ImagePlot summary visualizations of Desmet program, *left,* 1 and, *right,* 3. Full-color versions of these images can be accessed at http://www.create.human ities.uva.nl/#/. Images reproduced courtesy of EYE Filmmuseum/University of Amsterdam.

The image thus generated is zoomable. We find this type of visualization evocative in that it gives an at-a-glance impression of the color composition of a program, making it possible to see its shifting color schemes. In many respects, however, the visualization feels counterintuitive, for instance, because it places different segments on top of each other, thus obscuring their sequential relations. This is particularly evident when the user zooms in on the visualization, thus losing sight of the left and right edges of the composite picture. Also, montage visualization might not be the best fit for an analysis of the colors in a lengthy film program. The creation of such profiles is useful in particular for the analysis of color as a textual convention and as part of a close reading of individual films. Our goal, however, was to establish color palettes on a much larger scale.

In theory at least, summary visualizations as in Figure 6, although more abstract than montage visualizations and without an indication of the frames' sequential relation, are more productive when it comes to developing a cultural analytics approach to the study of early color. In the summary visualizations, every image of a sequence is layered onto the next, thus creating a canvas of colors with different qualities throughout the image. The predominance of one (median) hue (value), (median) saturation (value), or (median) brightness (value) in one area of the frame reflects a persistent occurrence of that hue, saturation, or brightness in that particular area of the frame throughout a program. The image thus allows us to understand, at least to a certain degree, the frequency of chromatic events in the frame, along the entire length of a program. In addition, comparison of film programs is possible, either by eyeballing it or with the help of additional ImagePlot functionality. For instance, by creating a scatter plot with

all summary images combined, one can arrange them onto a classic Cartesian grid, once again according to the same variables: in terms of hue and saturation, brightness and saturation, or hue and brightness.

In light of the many uncertainties concerning the color appearance of the (digital) films we worked with, we focused in our explorations on broad patterns in film hue. For the full Desmet programs, we can observe that the summary visualizations create predominantly light gray/greenish-colored pictures with some red sections; in program 3, in particular, a number of blue sections also appear (see the full-color versions of the black-and-white images in Figure 6). However, upon scrutinizing them, the user, much like in the case of the montage visualizations, has the urge to return to a closer inspection of the individual films to understand *why* some colors occur more frequently than others.

In addition to summary images of the entire programs, we also created visualizations of each of the films separately. Our intention here was to determine whether we would also be able to discern patterns according to genre. This gave intriguing, but once more inconclusive, results. Take, for example, program 1's first item, the nonfiction film *Target Practice of Atlantic Fleet US Navy* (Edison Manufacturing Company, 1912), the program's shortest film, with a duration of about seven minutes. Immediately noticeable in this visualization (see the color version of Figure 7, *left*) is the strong imprint of the contours of the intertitles and the overall brownish tone (due to the film's tinting). Tentatively, one might infer that this points to a prominence of text, associated in turn with the film's didactic nature. As a less conspicuous film in the program, its colors are more homogenous. The visualization of the Gaumont drama *L'Obsession du souvenir* (1913), featuring the popular actress Suzanne Grandais, in contrast, is the most spectacular one, at least from the perspective of color. The tinted, toned, tinted *and* toned, and stenciled sequences of this film merge into a very colorful picture (Figure 7, *right*). A visual check of both films does indeed reveal a relatively higher ratio of text and a much more consistent use of color in the first, thus confirming qualitatively a pattern observable after a series of quantitative operations.

Even if the pictures generated are ambiguous in what they show, there is one point they communicate rather unambiguously. As we mentioned, one of the most basic objectives of our color analysis experiment was to find out how we might be able to give users an at-a-glance impression of the sorts of color schemes audiences were exposed to when visiting cinemas showing Desmet's films. In our visualizations of three programs from the mid-1910s, any user can see that back then, watching black-and-white images was the exception rather than the norm. For indeed, while research efforts of the past two decades have made it abundantly clear, at least to experts, that silent cinema was
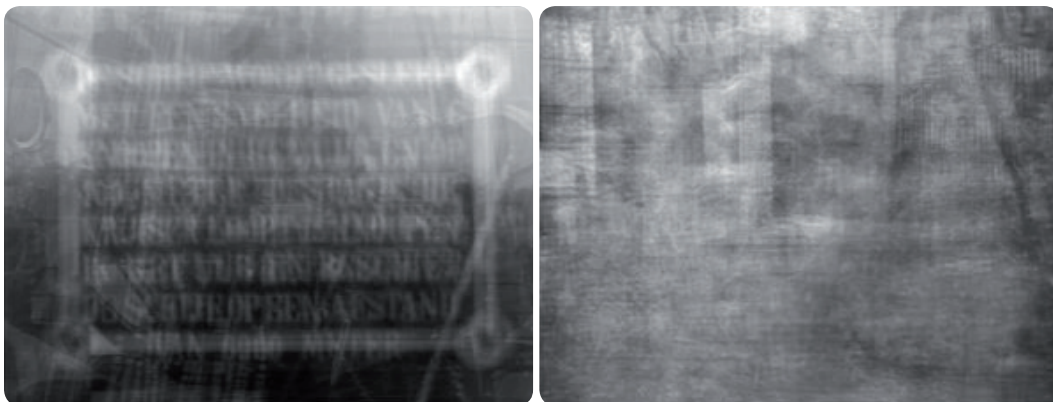
Figure 7. Black-and-white versions of ImagePlot summary visualizations of, *left, Target Practice of Atlantic Fleet US Navy* **(Edison Manufacturing Company, United States, 1912) and,** *right, L'Obsession du souvenir* **(Gaumont, France, 1913). Full-color versions of these images can be accessed at http://www.create.humanities.uva.nl/#/. Images reproduced courtesy of EYE Filmmuseum/ University of Amsterdam.**

overwhelmingly colored, it continues to be difficult to make this acceptable to both layman and scholars. The color visualizations directly confront viewers with the richness and range of color in films at the time and appeal to the public imagination much more powerfully than mere metadata do (the terms "color" or "applied color," or even "tinting" or "toning"). In this respect alone, one might argue, they serve a purpose.

However, further research and experimentation with color analysis tools can certainly help attain more ambitious goals and, specifically, help address more complex relationships between location, time, and the stylistic features of early films. For instance, by creating summary visualizations of full programs as well as individual films, analysis on two levels may be possible in the future. On one level, a comparison of summary images of programs or films might help film historians establish how the spectrum of film colors developed in the course of silent cinema's transitional years. It might enable them to detect changes in periods when attitudes toward color shifted and, with the help of additional sources, understand their relation to concurrent historical events. The limitation of the experimental visualizations here is that they combine the widely diverse color palettes of several films, resulting in single abstract images that do not necessarily do justice to the specificity of the patterns of the individual films. On the other level, a comparison of visualizations of individual films might also provide new insights into the use of color in various film genres or in specific genres over time.

Both types of envisioned research can be likened to the sort of "distant reading" that is promulgated by, among others, Franco Moretti.[37] However, the readings that we

propose here would be a version that, in line with Drucker's suggestions, emphasizes the contingency of our data and that therefore motivates historians to return to sources that can reveal concurrent shifts in production circumstances or audience attitudes. They would be armed, though, with new questions inspired by the patterns observed. This way, visualization software could function as a heuristic tool, helping researchers develop new perspectives on the materials they already work with.

CONCLUSIONS

For all parties involved in the "Data-Driven Film History" project, the development of and experimentation with digital humanities tools was a useful learning experience. Within the scope of an eight-month project, we focused on the methodological implications of working with digital research methods for both archivists and scholars rather than on the completion of a series of historical case studies. Clearly these should constitute the next step in our exploration of the Desmet data.

In designing the mapping interface, the archivists and researchers involved learned a great deal about coping with different data sets. Looking back on their experience as part of the project, EYE's archivists affirm that they gained more insight into (the quality of) their metadata. The exercise of combining different data sets also allowed them to add or correct information in their catalog. Additionally, the collaboration with external partners offered an opportunity for receiving feedback on both the quality of the data and the system's data export features. Overall, the project added archival knowledge of the Desmet Collection and its related data. The researchers, in turn, learned how crucial it is to understand the provenance, scope, and limitations of the data at their disposal and to make the discrepancies between them visible and transparent for others.

Our experiments in visualizing patterns in film color have proven that computerized analysis of large amounts of visual data, and their subsequent visualization, may reveal patterns that cannot be observed by simply viewing the films themselves. One lesson from our exercise was that the use of designated programs requires a great deal of computing power. Therefore it is crucial to determine early on the research questions one is seeking to answer. At the same time, visualizations can also reveal patterns that prompt questions other than those one started with.

For the Desmet demonstration tool specifically, one area to explore further is how we might productively integrate our mapping functionality with our color visualizations. Our main goal in trying out two types of visualizations simultaneously was to move on from the macropatterns of distribution networks to the micropatterns inside

the films that we know were shown—but more research is called for on how to bridge those two efforts. We could do this, for example, through a case study that explores whether any relations can be established between color and geographical distribution or between audiences and the palettes they were exposed to. Alternatively, we might try to find out whether there are any significant relations between color developments over time and geographical area.

The results we achieved as part of the project, however preliminary, prove once again how fruitful it can be for archivists and scholars to join forces. The archivists improved their metadata on the Desmet Collection, and the scholars learned from the EYE staff how to better understand their sources. This resulted in an original contribution to the development of scholarly digital methods: the conception of a tool that visualizes the uncertainties inherent to the data we work with.

Giovanna Fossati is chief curator at EYE Filmmuseum and professor of film heritage and digital film culture at the University of Amsterdam as well as the principal investigator of the project "Data-Driven Film History: A Demonstrator of EYE's Jean Desmet Collection." Fossati is the author of *From Grain to Pixel: The Archival Life of Film in Transition* (2009 and 2011) and coauthor, with Tom Gunning, Joshua Yumibe, and Jonathon Rosen, of *Fantasia of Color in Early Cinema* (2015).

Eef Masson is an assistant professor in media studies/preservation and presentation of the moving image at the University of Amsterdam. She is the author of *Watch and Learn: Rhetorical Devices in Classroom Films after 1940* (2012).

Julia Noordegraaf is a professor of digital heritage at the University of Amsterdam. She is the author of *Strategies of Display: Museum Presentation in Nineteenth- and Twentieth-Century Visual Culture* (2004/2012) and principal editor of *Preserving and Exhibiting Media Art: Challenges and Perspectives* (2013).

Christian Gosvig Olesen is a PhD candidate at the University of Amsterdam's Centre for Cultural Heritage and Identity. His research has been published in journals such as *NECSUS—European Journal of Media Studies* and *Cinema et Cie.*

Jasmijn Van Gorp is an assistant professor in television and digital culture at Utrecht University. She is an expert in the development and testing of digital tools for audiovisual heritage.

NOTES

1. In the United States, for example, the Library of Congress Motion Picture Conservation Center has digitized its famous Paper Print Collection of early motion pictures made before 1912. See "LoC Digitizing and Preserving the Nation's Motion Picture Legacy," http://www.mediageek.net/2009/03/loc -digitizing-and-preserving-the-nations-motion-picture-legacy/. In addition, the films that are included in the American Memory Online collection may be viewed at http://www.loc.gov/rr/mopic/ndlmps.html. Since the European Commission launched its Digital Agenda in 2010, many film heritage institutions in Europe also have begun digitizing parts of their collections. "Digital Agenda for Europe," http://ec.europa.eu/digital-agenda/en. Among those are films from and about World War I that have since been made available online via the European Film Gateway and Europeana. (This occurred within the context of EFG1914, a project that ran from 2012 to 2014, when 21 archives from 15 European countries and 5 other partners collaborated in the digitization of 761 hours of film and approximately 6,100 related documents. See the EFG1914 Project website, http://project.efg1914.eu.).

2. Mitchell Whitelaw, "Generous Interfaces for Digital Cultural Collections," *Digital Humanities Quarterly 9*, no. 1 (2015): 38, http://digitalhumanities .org/dhq/vol/9/1/000205/000205.html.

3. Compare Sarah Werner, "Where Material Book Culture Meets Digital Humanities," *Journal of Digital Humanities* 1, no. 3 (2012), http://journalof digitalhumanities.org/1-3/where-material-book-culture-meets-digital-human ities-by-sarah-werner/.

4. Whitelaw, "Generous Interfaces," 34.

5. The project was funded by the Netherlands Organization for Scientific Research (NWO) and ran from August 2014 to April 2015. For more information, see the project website, http://mappingdesmet.humanities.uva.nl.

6. Beside the authors, the project team consisted of Justin van Wees and Bart de Goede (for Dispectu, a software development company specializing in information integration and retrieval), Niels Sondervan, Jeroen Sondervan, and Dean Janssen (at Hiro, a firm that designs and develops online video applications for media, culture, and heritage institutions), Annelies van Nispen, Saskia Waterman, Rommy Albers, and Elif Rongen-Kaynakçi (all at EYE). The project also benefited from a user study conducted by Liliana Melgar (PhD student at Universidad Carlos III de Madrid and formerly an intern at EYE).

7. Another project member was Netherlands Institute for Sound and Vision, the national broadcast archive. Both institutions together digitized approximately twenty-two thousand hours' worth of moving images from their collections. For more information, see "Beelden van het verleden: 7 jaar beelden voor de toekomst" (Final report "Images for the Future" project, Nederlands Instituut voor Beeld en Geluid, EYE Filmmuseum, Nationaal Archief, Kennisland, 2015), http://beeldenvoordetoekomst.nl/publicatie/BVDT_eind publicatie_web.pdf.

8. It should be noted that, although "early cinema" typically refers to films of the period until the First World War, an ever growing body of work since the 1980s has focused on the pre-1907 period, arguing that its style was quite

distinct from that of later cinema, among others, in terms of narrativity (a reasoning which, since that time, has been challenged and ultimately refined a number of times). Considered from this perspective, the period during which Jean Desmet was active as a distributor (1907–16) can be designated as "transitional," because it still contains elements commonly associated with early cinema (e.g., its emphasis on spectacle or "attractions," the latter as defined in Tom Gunning's work), but also features characteristics of later cinema (e.g., certain kinds of narrativity). For this discussion, refer to the seminal papers by Tom Gunning, "The Cinema of Attraction: Early Film, Its Spectator, and the Avant-Garde," *Wide Angle* 8, no. 3/4 (1986): 63–70, and André Gaudreault, "Temporality and Narrativity in Early Cinema, 1895–1908," in *Film before Griffith*, ed. John Fell, 311–29 (Berkeley: University of California Press, 1983). Charles Musser exemplifies efforts since to refine this perspective. See his "Rethinking Early Cinema: Cinema of Attractions and Narrativity," *Yale Journal of Criticism* 7, no. 2 (1994): 203–32, and "A Cinema of Contemplations, a Cinema of Discernment: Spectatorship, Intertextuality, and Attractions in the 1890s," in *The Cinema of Attractions Reloaded*, ed. Wanda Strauven, 159–79 (Amsterdam: Amsterdam University Press, 2006). For a discussion on the characteristics of the films contained in the Desmet Collection, see Mark-Paul Meyer, "A Treasure Trove of Adventure and Experimentation," in *Jean Desmet's Dream Factory: The Adventurous Years of Film (1907–1916)*, ed. Marente Bloemheuvel, Jaap Guldemond, and Mark-Paul Meyer, 29–36 (Amsterdam: EYE Filmmuseum/Rotterdam: nai010, 2015).

9. "Eye and Desmet," https://www.eyefilm.nl/en/collection/selection/collection-dossiers/the-desmet-dossier/eye-and-desmet. For a complete list of films and posters, refer to Elif Rongen-Kaynakçi and Soeluh van den Berg, "Films and Posters in the Desmet Collection in the EYE Filmmuseum," in Bloemheuvel et al., *Jean Desmet's Dream Factory*, 161–89.

10. Ivo Blom, *Jean Desmet and the Early Dutch Film Trade* (Amsterdam: Amsterdam University Press, 2003), 334.

11. "The Desmet Dossier," https://www.eyefilm.nl/en/collection/selection/collection-dossiers/the-desmet-dossier; "Desmet Collection," http://www.unesco.org/new/en/communication-and-information/flagship-project-activities/memory-of-the-world/register/full-list-of-registered-heritage/registered-heritage-page-2/desmet-collection. Many of the films can be viewed online, via the EYE channel on YouTube, at https://www.youtube.com/playlist?list=PLB95B99C2420D32F3, or via the Open Images platform (http://openimages.eu/.en). The posters are accessible via EYE's page on the EFG website (http://www.europeanfilmgateway.eu/nl/node/1419).

12. Blom, *Jean Desmet*, 19–20.

13. Rixt Jonkman, "Any ID? Building a Database out of the Jean Desmet Archive," in *Networks of Entertainment: Early Film Distribution 1895–1915*, ed. Frank Kessler and Nanna Verhoeff (Eastleigh, U.K.: John Libbey, 2007), 312.

14. Anne Burdick, Johanna Drucker, Peter Lunenfeld, Todd Presner, and Jeffrey Schnapp, *Digital Humanities* (Cambridge, Mass.: MIT Press, 2012), 122.

15. David J. Staley, *Computers, Visualization, and History: How New Technology Will Transform Our Understanding of the Past* (Armonk, N.Y.: M. E. Sharpe, 2003), 34.

16. Ibid., 55 and 27, respectively.

17. Deb Verhoeven, Kate Bowles, and Colin Arrowsmith, "Mapping the Movies: Reflections on the Use of Geospatial Technologies for Historical Cinema Audience Research," in *Digital Tools in Media Studies*, ed. Michael Ross, Manfred Grauer, and Bernd Freisleben (Bielefeld: Transcript, 2009), 76.

18. See, e.g., Jeffrey Klenotic, "Putting Cinema History on the Map: Using GIS to Explore the Spatiality of Cinema," in *Explorations in New Cinema History: Approaches and Case Studies*, ed. Richard Maltby, Daniel Biltereyst, and Philippe Meers (Chichester, U.K.: Wiley-Blackwell, 2011), 58.

19. While the temporal dimension of this information was not the primary focus for the project, we did try to find ways around the sorts of issues with representing time that are mentioned by Verhoeven, Bowles, and Arrowsmith. For instance, in the mapping discussed in the following section, we decided to add a time slider that allows the user to restrict film distribution as represented on the map to specific time intervals. In addition, we included information on distribution and exhibition dates as attribute data in the title boxes. Both measures enable the user to see a film's distribution history in space and time. For further discussion of some of the ways in which time can be represented in mappings, see Ian N. Gregory and Paul S. Ell, *Historical GIS: Technologies, Methodologies and Scholarship* (Cambridge: Cambridge University Press, 2007).

20. Compare Johanna Drucker, *Graphesis: Visual Forms of Knowledge Production* (Cambridge, Mass.: Harvard University Press, 2014), 125.

21. Ibid.

22. Stephen Ramsay, "On Building," *Stephen Ramsay* (blog), January 11, 2011, http://stephenramsay.us/text/2011/01/11/on-building.

23. It should be noted that Collection EYE contains relatively more detailed information on Dutch films distributed by Desmet, even when no physical items are present in the collection.

24. All data sets can be accessed at http://www.create.humanities.uva.nl /results/datasets-and-coulour-visualizations-for-data-driven-film-history-a-demonstrator-of-eyes-jean-desmet-collection/.

25. During this process, it was possible to match 1,094 items, so that we ended up with 2,361 unique film titles. As the total amount of Jean Desmet films is estimated at fifteen hundred to two thousand, this means that there still are probably many duplicates (items that reference the same film but have different titles attached) within our final set of 2,361 films.

26. Drucker, *Graphesis*, 125–26.

27. To access this list, click on the EYE logo in the North Sea area of the map.

28. Refer, among others, to Daan Hertogs and Nico de Klerk, eds., *"Disorderly Order": Colours in Silent Film* (Amsterdam: Nederlands Filmmuseum, 1996); Monica Dall'Asta, Guglielmo Pescatore, and Leonardo Quaresima, eds., *Il Colore nel Cinema Muto* (Bologna: CLUEB, 1995); Giovanna Fossati, "When Cinema Was Coloured," in *Tutti i colori del mondo* [All the colors of the world], ed. Luciano Berriatúa et al., 121–32 (Reggio Emilia, Italy: Diabasis, 1998); Nicola Mazzanti, "Colours, Audiences, and (Dis)Continuity in the 'Cinema of the Second Period,'" *Film History* 21, no. 1 (2009): 69; and Joshua Yumibe, *Moving Color: Early Film, Mass Culture, Modernism* (New Bruns-

wick, N.J.: Rutgers University Press, 2012). For a discussion on the difficulty of assessing how many colored film prints circulated in the silent era, refer to Yumibe, *Moving Color,* 5–6.

29. Greg Miller, "Data from a Century of Cinema Reveals How Movies Have Evolved," *Wired,* August 9, 2014, http://www.wired.com/2014/09/cinema-is -evolving/. The Matlab software produces visualizations of the color palette of a given segment of a film on a *y*-axis and its temporal development and du- ration on an *x*-axis. See also James E. Cutting, "Current Research: Perception, Attention, and the Structure of Hollywood Film," http://people.psych .cornell.edu/~jec7/curresearch.htm (which, among other things, deals with his ongoing color research).

30. "Software for Digital Humanities," http://lab.softwarestudies.com/p /software-for-digital-humanities.html.

31. For the notion of "distant reading," see Franco Moretti, *Graphs, Maps, Trees: Abstract Models for Literary History* (London: Verso, 2007).

32. For help with reconstructing those, we are indebted to Rommy Albers, Maike Lasseur, and intern Leanne van Schijndel of EYE. One of the recon- structed programs may also be accessed online via https://www.eyefilm .nl/en/collection/the-desmet-dossier/cinema-owner/a-night-at-the-cinema -in-1915.

33. Mazzanti, "Colours, Audiences, and (Dis)Continuity," 78. One solution the author proposes in this context is to focus on Italian diva films, which tend to lean on conventional color schemes more prominently than certain other categories of films from the period. However, in light of our efforts to combine geomapping with color visualization, we chose instead a series of films shown in one location, to one (historical) audience. For the expression "disorderly ordered," see Hertogs and de Klerk, *Disorderly Order.*

34. For a discussion of the restoration of added color systems, refer to Giovanna Fossati, "New Approaches to the Restoration and Presentation of Colour in Silent Films," in *TECHNE Conserver l'art contemporain à l'ère de l'obsolescence technologique,* ed. Cécile Dazord and Marie-Hélène Breuil, 102–6 (Paris: Centre de recherche et de restauration des musées de France, 2013), and Daniela Currò, Claudy Op den Kamp, and Ulrich Ruedel, "To- wards a More Accurate Preservation of Color: Heritage, Research and the Film Restoration Laboratory," in *Color and the Moving Image,* ed. Simon Brown, Sarah Street, and Liz Watkins, 219–29 (New York: Routledge, 2013).

35. An interesting, recent series of examples is film scholar Kevin L. Fergu- son's studies of the Western genre, which uses the summary image function of ImageJ/ImagePlot to create "single frames of form and light" as the basis for evaluating his own color experience of Western classics such as *The Searchers* (John Ford, 1956) or *Johnny Guitar* (Nicholas Ray, 1954). See "What Does a Western Really Look Like?," https://medium.com/the-outtake/what -does-the-western-look-like-545981d93ae8.

36. Programs 1, 2, and 3, respectively, consist of 167,440, 135,182, and 159,705 images.

37. Moretti, *Graphs, Maps, Trees.*