„Kwartalnik Filmowy" no. 127 (2024)
ISSN: 0452-9502 (Print) ISSN: 2719-2725 (Online)
https://doi.org/10.36744/kf.2934

Miłosz Stelmach
Jagiellonian University
https://orcid.org/0000-0003-4004-7121

Cinema Counts: The Computational Turn and Quantitative Methods in Film Studies

Keywords:
quantitative methods;
digital humanities;
computational turn;
film studies;
data-driven humanities

Abstract
The aim of this text is a critical analysis of current developments and potential applications of quantitative methods in film studies. Within its scope, a concise reconstruction of the methodological foundations, historical development, and key achievements of statistical, experimental, and digital humanities tools in relation to audiovisual media research is conducted. This involves a review of the phenomena that have developed so far as well as a philosophical consideration of the sources, consequences, and potential limitations of quantitative thinking in an area traditionally occupied by the humanities. Quantitative methods are not considered here as a replacement for existing paradigms, but rather as their complement, extension, and often inspiration. This allows to understand the current transformations but also integrate them with traditional research approaches, and identify the pitfalls and difficulties associated with this paradigm shift.

In an era where film predominantly takes the form of a digital file, filmmaking itself largely depends on the use of elaborate software, and human creativity in this process is supported – and sometimes even replaced – by artificial intelligence, our methods of understanding and describing cinema are also undergoing significant changes. Just as cinema has become digital in recent decades and films have become data sets, analyzing them increasingly involves performing operations on data. As a result, we observe a growing trend toward studying culture with the aid of computational methods, primarily those borrowed from the exact sciences (computer science, statistics) and social sciences (sociology, economics, cognitive science, communication science). The aim of this article is a critical analysis of their current developments and potential applications in film studies.

The article encompasses a concise reconstruction of the methodological foundations, historical development, and key achievements of statistical, experimental, and digital humanities tools in relation to audiovisual media research. This involves a review of the trends that have developed so far and the most significant ongoing research projects and scientific reflections in this sphere, as well as a philosophical consideration of the sources, consequences, and potential limitations of quantitative thinking in an area traditionally occupied by the humanities. Importantly, quantitative methods do not serve here as a counterbalance or replacement for existing paradigms in film studies, but rather as their complement, extension, and often inspiration. This, in turn, will allow us not only to understand the transformations currently taking place, but also to determine how new digital and computer tools can change the way of practicing humanities in the future, to identify the best way to integrate them with traditional research approaches, and to see how we can avoid the pitfalls and difficulties associated with the paradigm shift occurring in recent years.

The computational turn and the rise of digital humanities

The need to address this issue arises from recognizing the significance of widely discussed phenomena, such as the so-called computational turn in contemporary humanities[1] and the emergence of data-driven humanities.[2] They both denote the increasing reliance of fields like literary studies, cultural studies, linguistics, art history, and others on large data corpora and digital tools for their collection, cataloging, analysis, and presentation as a natural consequence of technological development, particularly the widespread availability of powerful processors, user-friendly data acquisition and analysis software, and the vast amounts of data generated and made accessible through the digitization of cultural resources and Internet access. As a result, digital humanities tools, big data, and AI mechanisms have become integral to numerous scientific projects, supporting or replacing traditional methods of acquiring and transmitting knowledge to varying degrees. These tools help address existing questions and generate new types of problems and research approaches previously beyond the scope of many humanities disciplines. Consequently, as David Berry noted over a decade ago, *computational technology has become the very condition of possibility required in order to think about many of the questions raised in the humanities today*,[3] and his words seem increasingly true each year.

Although quantitative methods can have diverse applications and manifestations in the humanities, the clearest sign of this reorientation is the shift from a detailed analysis of individual cultural objects and texts to what Matthew Jockers calls *macroanalysis*[4] and Franco Moretti terms *distant reading*[5] in literary research. This primarily signifies a change in the scale of analysis, which increasingly addresses broad trends – such as large historical patterns, extensive collections of data and metadata, empirical data from experiments, or multifaceted analysis of textual or visual corpora. Working with appropriately operationalized data allows scholars to conduct analysis and draw conclusions based on thousands or even millions of cultural objects (words, images, numbers, or other types of information), which time and cognitive limitations make inaccessible to individual researchers or even entire teams working traditionally. These data, often collected in databases – frequently created in a fully or partially automated way – are analyzed using statistical tools or mathematical models and presented through algorithmically generated graphs, charts, or visualizations, providing access to new kinds of knowledge about culture. Lev Manovich advocates elevating such activities to a distinct subdiscipline of contemporary humanities, calling it *cultural analytics* and defining his and his team's efforts as *the use of computational and design methods – including data visualization, media and interaction design, statistics, and machine learning – for exploration and analysis of contemporary culture at scale*.[6]

These processes inevitably relate to the idea and practice of digital humanities, which has been developing over the past few decades, and which occupies the intersection of humanities and computer science or digital technologies. It is not synonymous with the application of quantitative methods in the humanities, as the latter saw regular use already in the pre-digital era – though to a much lesser extent than today – and can still employ relatively simple mathematical or graphic tools that one might not consider digital humanities. Also the other way around, just as only a portion of quantitative research fits within the broadly understood digital humanities, the latter can assume various meanings, only some of which involve computational methods. According to the typology introduced by Camille Roth, digital humanities might include three different but overlapping tendencies: digitized humanities (*constitution, management, and processing of digitized archives*), numerical humanities (*emphasizing mathematical abstraction and the development of numerical and formal models*), and humanities of the digital (*study of computer-mediated interactions and online communities*).[7]

Among these three areas, only the second – numerical humanities – aligns directly with the tendency to apply quantitative methods in cultural research, although the latter often also utilize resources created through digitized humanities. After all, the computational turn in contemporary humanities would remain impossible without the datafication of society and culture,[8] i.e., the digitization and conversion of cultural texts and information about us and the world we live in into data. The combination of such datafication of culture and its products with the scientific approach characterized – at least in its strict version – by formulating and testing hypotheses verifiable using quantifiable measures and facts creates a new way of thinking. Researchers guided by this approach treat issues considered within the humanities as empirical problems, which one can resolve or at

least operationalize and quantify using available data. Thereby, they illustrate the famous aphorism of astrophysicist (and novelist) Carl Sagan: *Science is a way of thinking rather than a body of knowledge*. And this way of thinking has occasionally manifested itself in the history of film studies and is now beginning to grow into a significant part of it.

All the aforementioned changes are also evident in the broadly defined study of film, media, and audiovisual culture, which incorporates many tools developed in recent years – on the one hand, in the field of literary studies or art history, and on the other, in social sciences such as economics, sociology, or psychology, all of them complemented by new methods and proposals. Although their diversity, scale, level of complexity, and scope continue to grow, it seems that one can distinguish certain main tendencies and research directions along with accompanying methods. Therefore, I would like to propose a typology that encompasses three basic types of applying quantitative methods in cinema studies, defined at the same time by the methods used, academic and institutional genealogy, and the most frequent subject of research. The typology remains highly simplified, and the boundaries between those categories inevitably blur at times, but I believe that it allows for a useful differentiation and operationalization of the basic vectors of current research at the intersection of film studies and computational methods.

Film metadata analysis: Studying the context

The first major area of intersection between film studies and quantitative methods concerns fields where various types of data, especially numerical data, naturally appear, and where the computational possibilities offered by digital technology simply make data collection and analysis easier and more adequate. These are especially areas related to the economics and sociology of cinema, as well as production studies, which may include information on the height of production budgets and funding sources, revenues of titles in various distribution channels, audience demographic, or characteristics of national film industries, including the number of movies produced each year. Such information can be termed metadata, i.e., 'data about data,' as it concerns a set of various accompanying details, phenomena, or objects, not the films themselves.

These issues had been of interest to scholars way before film studies emerged as a separate academic discipline. Since the foundational study *Zur Soziologie des Kino*, a 1914 doctoral dissertation by Emilie Altenloh,[9] sociologists focused on audience statistics,[10] while economists started analyzing the financial side of the film industry,[11] and this line of research remained dominant for several decades.[12] However, access to those and similar data is much simpler now than ever before, as are the possibilities for processing and presenting the results. Consequently, the catalog of metadata that researchers consider has expanded, including slightly less obvious areas and conclusions. Such metadata may include film ratings and reviews posted on websites or digitized editions of the press, information about the cast and crew, cinema programs, statistics on queries for

specific content in search engines, or the titles of the works themselves. Furthermore, new methods of processing and presenting these data often result in visually appealing but also more comprehensible forms, such as various types of charts, visualizations, or maps. The latter are part of a broader phenomenon sometimes referred to as a spatial turn in the humanities and social sciences – excellently illustrated, for example, by the long-term research conducted by Deb Verhoeven and her collaborators as part of the Kinomatics project, presenting data on the transformation and outcomes of Australian cinema exhibitors,[13] global flows and behaviors of international blockbuster audiences, or film cultures in different cities.[14] Another related undertaking is the Film Circulation project, which uses quantitative methods to analyze the complex network relations of the film festival sector using programming data.[15]

As a result, contemporary research often employs models allowing, for example, the analysis of economic data to determine the profitability of film projects,[16] estimation of the size of theater audiences in times when detailed statistics were not collected, and comparisons of the relative popularity of film works, as John Sedgwick did with his POPSTAT and RelPOP formulas.[17] The second fundamental trend is using metadata to reconstruct historical macrotrends encompassing a chosen aspect of cinema, such as changes in the size of film crews and the increasing specialization of professions on set,[18] the frequency and nature of remakes in the American production system,[19] the growing diversification of European audience film preferences in the first decades of the 20th century,[20] or the gender division of labor in film production in the 21st century as measured by the Gender Equity Policy Analysis project.[21] This strand is sometimes called digital film historiography,[22] and its genealogy closely relates to the approaches associated with so-called New Film History.[23] Finally, this field also encompasses the sociology of film audience and reception studies, which empirically approach the circulation and reception of film works or other cultural texts,[24] and which constitute a rich separate subdiscipline, historically related mostly to the media and communication studies.

Psychology, neurosciences, and biosciences: Studying the viewer

The second key area of film studies research utilizing quantitative methods for years is the perspective of body and cognition studies. Researchers with experience and tools specific to disciplines such as psychology, biology, or neurocognition employ them to empirically address questions about what happens to the viewer when interacting with a film. Thus, they measure the viewers' psychomotor conditions, physiological reactions, cognitive activities, beliefs, and reflexes, as well as the influence of cinema on the audience – very often using experimental methods and the statistical analysis of results obtained that way. This area also includes historical precedents dating back to the 1930s and 1940s, mostly focusing on behavioral psychology and its attempts to address the question of cinema's impact on the viewers' behaviors and beliefs.[25] Studies of this type frequently concerned children and took place in the atmosphere of moral panic resulting from

the belief in the new medium's potential harmfulness for young audiences, such as the famous Payne Fund Studies, conducted in 1929-1932.[26]

Contemporary researchers more often employ advanced equipment not created specifically for film studies research, and use it to monitor aspects such as reaction time to stimuli appearing on the screen[27] or the brain activity of people watching films,[28] and various other biological processes, such as exhaled air composition[29] or the bloodcurdling effects of horror movies.[30] Especially the neurocognitive approach has seen extensive development, going beyond single case studies and aiming toward systematizing the entire subdiscipline, whose proposed names include neurophenomenology of film experience,[31] NeuroCinematics,[32] psychocinematics,[33] or neurofilmology.[34] All of them treat the film image (and sound) as a physical and cognitive stimulus, seeking to measure, describe, and interpret the human organism's reaction to it – be it physical, psychological, or emotional. For example, researchers have examined the emotional response by checking the correlation between understanding causality in film narratives and personality traits.[35]

Also, numerous eye-tracking studies have grown into a separate vein of research, involving the recording of the viewers' eye movements to track how they direct their gaze – namely which screen elements they focus on at a time and how long the focus lasts. Such studies help reveal which aspects of a film work (e.g., motion, framing, exposure, camera work) attract the viewers' attention and to what extent they do so, how much we differ from each other in what we look at, how the visual presence of subtitles affects our visual attention, or how our visual activity changes with age.[36] Although focused on the visual layer for obvious reasons, eye-tracking studies do not necessarily ignore the auditory sphere, as shown by researchers seeking connections between acoustic effects (sound, music, sound equipment quality) and the viewers' visual activity.[37]

Due to the high level of specialization, both in terms of scientific knowledge and purely technical aspects related to operating advanced devices – such as eye-tracking scanners or, especially, the MRI and fMRI equipment – these studies are still mostly conducted by scholars with bio- and neurocognitive background, rarely by film scholars or with their involvement. This has consequences for designing experiments and posing research questions, which more often concern the mechanisms of human cognition in contact with film fiction rather than, for example, the characteristics of films in this context, and for scientific communication – the way of developing and presenting results or their publication venues.

Cinemetrics and distant viewing: Studying the film

Finally, the third branch of quantitative studies in the realm of cinema, most closely associated with the films themselves, attempts to quantify its content and stylistic elements, measuring various parameters or significant components. This area also has fairly distant historical precedents, in which people like Dorothy B. Jones or Jacques Doniol-Valcroze noted the occurrence of certain tropes, character types, or narrative structures in selected sets of films and created simple statistics

Riders of Justice, dir. Anders Thomas Jensen (2020)

regarding their prevalence.[38] The first person to attempt this approach with greater scholarly rigor was Barry Salt (1970s), who focused on the frequency of editing cuts and historical changes in this area.[39] It was largely through the inspiration drawn from his work that the entire tradition associated with the Cinemetrics project, initiated in 2005, developed in the 21st century. Scholars gathered around the website established and maintained by Yuri Tsivian at https://cinemetrics.uchicago.edu (formerly https://cinemetrics.lv) – including Barry Salt, Mike Baxter, Nick Redfern, and others – developed the theory and practice of statistical analysis of film style, initially focusing primarily on measuring shot lengths, transitions between them, shot sizes, numbers of characters, and so on.[40] Measurements were mostly manual, with only slight technological assistance provided by the simple Cinemetrics Measurement Tool software, thus limiting the scope of their analysis. Nevertheless, they laid the groundwork for more advanced contemporary studies of this kind. Due to these contributions and their similarity to the concept of stylometry adopted in literary and linguistic studies,[41] I will use the term 'cinemetrics' in an extended sense: as a concept denoting any research aiming at a quantitative approach to the film work itself – its formal, structural, or narrative aspects.

In the meantime, work on algorithmic tools has continued since the turn of the century, allowing many measurements to proceed automatically or semi-automatically.[42] In recent years, computer vision technology – utilizing methods of artificial intelligence and machine learning – has developed rapidly, enabling automated 'watching' of audiovisual works and thereby providing materials for large-scale analysis of their features and aspects. These methods allow for the recognition of objects appearing in the frame, its composition, or formal features such as brightness, color saturation, or lighting gradients in large (audio)visual corpora. Taylor Arnold and Lauren Tilton have taken advantage of this potential expansion of the research scale, proposing the concept of *distant viewing*, understood as *a methodological and theoretical framework for studying large collections of visual material*.[43] They even developed their own software tool, the Distant Viewing Toolkit, freely available online for use by other researchers. Arnold and Tilton employed it, among other things, to compare the narrative structures of American TV shows from the 1960s and the 1970s, or to identify genre codes on movie posters.[44] Another tool frequently used by quantitatively oriented researchers in audiovisual culture is the widely available You Only Look Once (YOLO) package, enabling the recognition of various elements in visual materials. Although not created for film analysis purposes, it can be trained for the specific needs of film studies.

These and other available tools enable precise and data-driven analysis of properties such as color composition and saturation,[45] lighting and brightness, camera movements and the scale of individual shots (studied for some years now by the CineScale project),[46] or the musical parameters of the soundtrack.[47] Although occasionally related to studies conducted in literature (e.g., the regular use of stylometry tools for analyzing film dialogues[48] or screenplays[49]) or art history and visual culture analysis (research on visual material corpora in projects like Cultural Analytics), this area seems to utilize the tools most specific to film studies, taking into account the specificity of cinema as a medium and as a formal system, created using parameters such as image texture, framing, motion, or the passage of time.

The three areas outlined above do not exhaust the possibilities of using quantitative methods in film studies, but I reckon that they delineate relatively separate, distinct fields that have flourished in recent years. Each of them attempts to answer slightly different questions traditionally posed within the scope of film studies but employs different methods for this purpose. Certainly, all of them will evolve, influence each other, and intersect, perhaps leading to the emergence of entirely new paths and research directions. Meanwhile, to understand and appreciate the newly outlined possibilities facing contemporary humanities, but also to subject them to critical analysis, one should examine the potential methodological, philosophical, and sociological consequences of these changes. Such is the focus of the second part of this article.

Lucky numbers, or the benefits

If quantitative research in film studies is to be more than a novelty on the fringes of main research traditions, it must address the discipline's vital questions and preoccupations as well as introduce new elements and benefits. Therefore, one should highlight the most important gains pointed out by scholars involved in computational research in culture.

The first gain entails changing the type of questions asked. Quantitative methods serve not only to help resolve the already known questions more precisely and confidently, but also – in some situations – to transform their very nature. One imperative is to formulate research problems in a way that enables one to address them using empirical data and verifiable hypotheses. David Bordwell argues for the inherently empirical nature of film studies, accurately noting: *Film studies is an empirical discipline. Film scholars attempt to describe, analyze, and explain artifacts and events pertaining to the medium of cinema. These artifacts and events existed, and so they're amenable to rational-empirical investigation*.[50] According to this definition, almost any inquiry that is not an abstract theory or philosophy of cinema but contains a historical, interpretative, or analytical element should indeed be called empirical, as it has roots in the matter of the films themselves. The aim of quantitative research, however, is to transform this empirical nature into operationalizable categories that one can measure and interpret.

A key role in this belongs to the change in scale – that is, greater focus on macrotrends and/or large-scale historical processes to which individual researchers had very limited access using traditional methods. Arguing for his turn toward cultural analytics, Lev Manovich writes: *My original motivation for turning to computational methods and big data came from the realization that the scale of culture in the twenty-first century makes it impossible to see it with existing methods*.[51] According to various estimates, global cinematography produces from nine thousand to even sixteen thousand feature films annually – along with unknown, enormous amounts of documentaries, short films, and series – which clearly makes it impossible to familiarize oneself with all of them. In practice, researchers usually focus primarily on works or creators that are already recognizable or considered significant for some reason, posing specific rather than more general questions. Working with big data and automated data processing allows for an inclusion of

information beyond the capabilities of an individual or even an extensive research team. This may help avoid the 'great slaughterhouse of cinema,' analogous to the *slaughterhouse of literature* described by Franco Moretti, in which most books ever written are 'slaughtered' – that is, made invisible to literary scholars because they are practically absent from studies, analyses, or reviews.[52] Until now, film scholars have also been blind to most films ever made, formulating their opinions and hypotheses primarily based on a handful of films that have achieved a sufficiently prominent position in culture. While this position is worth investigating in itself because it illustrates the mechanisms of creating cultural hierarchies, the situation where we draw far-reaching conclusions about the nature and transformations of the film medium based on a tiny fraction of its manifestations calls for deep rethinking. Quantitative methods offer mechanisms for this reconsideration.

The second gain entails supporting interdisciplinary and collaborative work. The necessity of drawing from various disciplines (e.g., economics, sociology, linguistics, psychology) and utilizing technical skills (e.g., computer science, statistics) for acquiring, processing, or visualizing data forces researchers interested in quantitative methods to acquire diverse competencies beyond the scope of a single discipline. This particularly promotes inter- and transdisciplinary collaboration, where experts from different backgrounds can bring their unique perspectives to the project. It fosters the practice known in many scientific disciplines, which build teams or laboratories able to tackle challenges that one researcher could not handle alone, and to do so in a more multifaceted way. It also integrates film knowledge with other findings, such as those related to human perception, cognitive psychology, or social mechanisms, allowing for a more holistic, cross-disciplinary reflection on contemporary audiovisual culture and its significance for individuals and societies.

Current experiences in this field show that such is the exact direction in which research is moving: unlike in traditional humanities, single-author works are rare in digital humanities journals and conferences. Instead, works often gather researchers from different disciplines, departments, universities, and sometimes even countries. Nothing fosters the development and quality of scientific work more than the diversity of views and approaches, mutual inspirations, and the testing and validation of methods and integration of different areas of knowledge. This breaks the traditional shortcomings of advanced scholarly work in the humanities – the high degree of theoretical dogmatism and hermeticism of resulting works. Quantitative research, though requiring a higher entry level (e.g., statistical literacy), makes research more communicable and translatable into other areas of knowledge about culture and society.

The third gain entails the increasing intersubjectivity and verifiability of research. Allowing for a caricatural oversimplification, one might say that a film historian studying the development of a particular stylistic trend in contemporary cinema based on films they have watched resembles a climatologist trying to trace global warming using measurements taken at home with a household thermometer. Their measurements may align with global trends, but it is easy to question whether a handful of arbitrarily selected, somewhat random samples accurately reflects the climate changes worldwide. Similarly, in cinema studies,

when attempting to trace changes in, for example, the use of long lenses over recent decades, or other stylistic development, an individual can watch and analyze perhaps dozens or, at most, hundreds of films. Of course, traditional historiographical methods remain more refined than random analysis, drawing from deep understanding of film historical processes and vast knowledge of the field, but questions about size and selection criteria of the analyzed corpus still arise. Surely, this is by no means a new problem; there have been decade-long disputes on the matters of sampling, statistical significance of observed effects, and strengths and weaknesses of quantitative versus qualitative methods – especially in social sciences, where both approaches are equally prominent and considered legitimate. This debate and the rich literature it has spawned needs confrontation with the current debates on digital humanities to avoid re-inventing the wheel over and over because of disciplinary divides. In any case, quantitative methods can overcome at least some of the potential difficulties described above: on the one hand, by considering a much larger dataset using digital tools, and on the other, by drawing on extensive traditions in other disciplines regarding the selection of research samples and establishing the significance of the measurements.

The result should be increased verifiability and intersubjectivity in the humanities by reducing the arbitrariness of material selection for analysis and of the conclusions drawn, confronting theoretical judgments and entrenched opinions with hard data, and identifying historical heuristics and generalizations associated with researchers' natural cognitive errors and biases. Utilizing statistics or empirical and experimental data can serve as a means of reexamining the claims previously made in the literature and the diagnoses based on qualitative analyses, theoretical suppositions, and researchers' experience-based intuition. Such verification endeavors shape new methodological standards concerning, for example, the selection of studied works, the argumentation, and hypotheses presented in a way that – at least in some cases – allows for their factual resolution.

Limitations, problems, pitfalls

The benefits of the quantitative approach in film studies find indirect proof in the dozens of published works of this type, including those cited earlier, whose authors convincingly and productively analyze culture – including film culture. However, this does not mean that we should uncritically accept new paradigms as unquestionably and obviously better than existing ones, let alone free from errors. Intensive work in this area over recent years reveals not only advantages and applications but also limitations, problems, and pitfalls. These are not merely methodological gaps in selected methods or mistakes made by individual researchers on specific topics, but broader recurring patterns. Similarly to the benefits, these difficulties permit categorization into three most frequently occurring categories.

The first category is the illusion of objectivity and naive scientism. More specifically, a common concern raised by skeptics is the illusion of objectivity stemming from excessive faith in the unequivocal character of quantitative data due to its mathematical nature associated with scientific undeniability and certainty. The most striking manifestation of this thinking was the views of some re-

searchers, who predicted the 'end of theory' at the early stages of big data analysis. In its most crude version, this stance posited that the vast amount of data could resolve any issue that had previously been the subject of 'merely' theoretical – and thus implicitly speculative and unscientific – disputes and inquiries.[53] Today, few express this view in such a radical form. Most researchers realize that the opposite is true: while data themselves can indeed be objective, they always appear in a particular context that is not. It depends on who defines, acquires, categorizes, interprets, and uses the data. This is why, as the very titles of significant publications on the subject suggest, *there is no such thing as raw data*,[54] and the phrase *"raw data" is an oxymoron*[55] because the data available to us always result from some human intervention. For instance, the ignored biases or selectivity of the input data may become invisible at the stage of results and their discussions later in the research, and hence in subsequent works using them. As a result, databases, programs, and technologies created by humans often reproduce human cognitive errors, such as racial biases, albeit in less visible ways.[56]

Recognizing the issue of so-called raw data necessitates a critique of the belief in the end of theory and suggests that the outcome of developing digital and quantitative methods in the humanities should be quite the opposite. New possibilities and the excess of data they create require new theories and carefully considered methodologies to guide them and find the most appropriate tasks or questions for them. The call for a return to theory in quantitative research is complemented and fulfilled in related calls for their politicization, even in such hard research as eye-tracking studies. William Brown comments on this area as follows: *To claim that we can isolate films and film viewing from a human world that is perhaps always political, and to claim that we can then analyze films 'as they are,' is perhaps absurd: films 'as they are' are part of a political world*.[57] This also creates a challenge for work on archival materials, recently scrutinized in the context of film studies,[58] which need updating in light of new digital ways of handling archival materials and datasets based on them.

The second category is the mechanical flood of data. Specifically, the belief in the inherent value of data can also lead to simply generating it mechanically, equating this process with scientific activity. This may result in an overabundance of data sets, databases, charts, and numbers that do not necessarily provide a better understanding of a topic or a solution to the challenge at hand. Byung-Chul Han identifies this issue as he notes: *Correlation is the most primitive form of knowledge, not even capable of ascertaining the relationship between cause and effect. It is so. The question of why becomes irrelevant; thus nothing is understood. But knowledge is understanding. Hence Big Data renders thought superfluous*.[59] As more data accumulates, expert knowledge to select the most relevant topics, problems, and results, as well as to synthesize subsequent discoveries and reintegrate them with the existing body of knowledge, will become more critical than ever.

We must also be cautious of the fact that when we uncritically equate the production of knowledge with data generation, we often mechanically reproduce our limitations and perspectives, simply applying our cognitive biases on a larger scale. As Cathy O'Neil has shown, algorithms tend to replicate societal biases, such as racial prejudices, due to the human factor involved in their design and the ways

of collecting and interpreting data.[60] This issue aligns with the earlier call for the politicization of quantitative research, but even if we set the social consequences aside for a moment, a pure research perspective also warrants particular attention. Contrary to the declared primary advantage of encompassing cultural objects that might otherwise end up in the slaughterhouse of literature, quantitative research is also prone to biases, such as the availability bias and the popularity bias. These arise from the fact that research often utilizes the most popular works, in our case films, but this also applies to studies of music, literature, websites, and so on. This happens for several reasons, such as implicit (or explicit) assumptions that these works are the most representative or that, as the most popular ones, they best reflect the viewers' experiences. In addition, they are the most accessible and more likely to be digitized, and have the most (meta)data available. The fact that film scholars rarely conduct such studies means that the selection of films is less informed by the field's internal needs and developments. Thus, paradoxically, we may repeat our cognitive errors and selective content choices on a larger scale, replacing the manual slaughterhouse of literature with its newer version – digital, automated, and statistically significant.

The third category is redundancy and excessive scientific apparatus. In one opening scene of the film *Riders of Justice* (dir. Anders Thomas Jensen, 2020), statisticians at a large corporation present the results of their work to the company's internal committee. A slightly shortened version of the dialogue goes as follows:

– The algorithm tracked 82,504 registration plates from 46 municipalities and linked them to tax refunds to generate statistical data. The algorithm indicated that the lowest-income group drives Kia, Fiat, and Hyundai. The middle-income group drives larger Toyota models, Ford, and Volvo. The highest-income group primarily chooses Mercedes, Tesla, and Audi. How we calculated this...

– I'll interrupt you, Otto.

– Yes?

– How much time did you and your team spend creating this algorithm?

– 46 weeks. But mostly at night.

– So we spent a year and a fortune on an algorithm that tells us that poorer people drive Kias and richer people drive Mercedes?

After presenting the results, the characters are immediately dismissed, and this humorous scene illustrates the risk of quantitative research in traditional humanities: sometimes long data sheets, complicated formulas, and meticulously compiled columns of code lead to conclusions that are obvious to anyone who has even casually studied the subject. The third danger, therefore, is redundancy and excessive scientific apparatus, which means using complicated tools to solve problems that traditional methods handled quite well. This can cause situations where substantial resources, time, and researcher attention are devoted to complex work on issues that insightful researchers have already described accurately using qualitative analysis tools. One example is eye-tracking research, where an analysis of viewers watching a scene from the film *There Will Be Blood* (dir. Paul Thomas Anderson, 2007) revealed that their attention focused primarily on movement, color, and events in the center of the frame, especially if they were in the foreground, sharp, and included human faces and hands[61] – basically the same

There Will Be Blood, dir. Paul Thomas Anderson (2007)

conclusions as those reached earlier by David Bordwell through a traditional analysis informed by his cognitive film theory.[62]

On the other hand, such cases can serve as examples of the complementary character of qualitative and quantitative research methods, and triangulation via textual analysis and viewer analysis that encourages desirable attempts to empirically verify the hypotheses posed by theorists. In this light, scientific confirmation of existing views represents a natural path of research work – from hypothesis to its subsequent, increasingly robust evidence. This is also a response to the often-raised need for replicability of results in the humanities, which have so far often evaded close scrutiny.[63] These potential benefits do not entirely eliminate the risk that, in the worst case, elaborate statistical apparatus or digital humanities tools may be used to obscure the banality of some studies and their conclusions, using costly and time-consuming research and complex expressions to convey intuitively obvious facts.

Having listed these potential pitfalls that can affect numerous quantitative studies, one must acknowledge that most researchers in the field are aware of these problems and challenges, and regularly point out the limitations of their own work and its possible conclusions. However, their presence should not halt inquiries in this area but rather make us aware of their often auxiliary and supportive relation to traditional research competencies. One pioneer of the stylometric method in literature, Jan Rybicki, likes to use the metaphor of an X-ray, comparing it to statistical text analysis. An X-ray is a better and more reliable method of understanding what is happening inside a patient's body than traditional auscultation – using technology provides greater accuracy and reveals things impossible to see otherwise. However, it does not deprive the doctor of agency and control: it remains up to them to interpret the X-ray, make a comprehensive diagnosis, and propose further examinations or treatment. Similarly, quantitative research is not intended to mechanically and unequivocally explain all aspects of cultural activities and their results or resolve all the theoretical disputes. Instead, its goal is to provide more thoroughly verified information and, consequently, perhaps more precise concepts and better arguments to continue these disputes.

A new episteme and computational imagination

It is also essential to emphasize that the computational turn in contemporary humanities and film studies is not merely an addition to existing research tools or the introduction of new methods. Unlike previous decades, where successive theories and methodological orientations frequently expanded the landscape of humanistic inquiries without altering its general character, we are witnessing a profound reorientation of the entire culture and idea of scholarly work, as well as the very notion of knowledge in the studies of culture, humanity, and society. This change occurs at the epistemological level, encompassing the methods of understanding the world, and perhaps even at the ontological level, concerning the very nature and functioning of knowledge, as well as our perception of reality and access to knowledge.

David Berry argues that in the 21st century, knowledge is being transformed into information, which makes computer science a foundational basis for all scientific research, influencing other disciplines too. *A computer requires that everything is transformed from the continuous flow of our everyday reality into a grid of numbers that can be stored as a representation of reality which can then be manipulated using algorithms. These subtractive methods of understanding reality ("episteme") produce new knowledge and methods for the control of reality ("techne")*.[64] Viewing reality as a stream of data and scientific research as the skill of analyzing, manipulating, and recognizing patterns within that stream creates a new kind of researcher, and perhaps even a new type of subject, which Berry calls a *computational or data-centric subject*.[65]

This subject is a product of the digital age, with its numerical approach to the world and knowledge about it. However, the philosophical roots of this attitude are much older. We can trace its manifestations across different historical epochs, starting with the ancient Greek Pythagorean formula "All things are numbers." Alfred W. Crosby argues that the end of the Middle Ages and the Renaissance in Europe brought a fundamental change in the approach to reality, intensifying the quantification attempts, especially regarding parameters such as time and space.[66] The culmination of this process appears in Gottfried Wilhelm Leibniz's call *Calculemus!* (*Let us calculate!*), which Jonathan Gray posits as an expression of *computational imagination* – the belief in the numerical nature of reality, which one can best understand through the exact sciences.[67] This belief became the foundation of (post-)Enlightenment science, and even though the humanities remained impervious to it for a long time, one notices its influence already in the earliest attempts at quantitative analysis of culture, including cinema. The aforementioned pioneer in this field, Dorothy B. Jones, described her work in 1942 as an *experimental study designed to provide an instrument capable of measuring with scientific exactness the content of each motion picture as it is released*.[68] However naive such declarations about the 'scientific exactness' of film content analysis may sound today, they express the evolving contemporary rendition of the above-defined 'computational imagination.' In the 21st century, they have taken their most extreme form in the diagnosis of – and sometimes even the demand for – the 'end of theory' as described above.

The consequences of this paradigmatic shift exceed the realm of ideas and epistemological-ontological foundations of science, and extend to its concrete social and economic dimensions. One cannot consider the computational turn in isolation from the sociology of science, a discipline that explains the ways of producing and legitimizing knowledge, influenced by changing beliefs about the role and credibility of scientific research over time. Referring to C. P. Snow's canonical diagnosis of *two cultures* that inform modern intellectual life and reflect the divide between the humanities and the sciences,[69] one can state that for many decades, film studies as a young discipline focused on a culturally less prestigious medium like cinema but naturally gravitated toward the studies of art, theater, and literature, seeking academic legitimacy through a humanistic and hermeneutic approach. However, those vectors have now reversed: film scholars no longer need to prove the artistic or cultural significance of their field, whereas the humanities as a whole are under scrutiny compared to hard sciences – valued for their sup-

posed objectivity, verifiability, and professionalism. This view is encapsulated in the famous aphorism attributed to one of the fathers of modern physics, Ernest Rutherford: *That which is not measurable is not science. That which is not physics is stamp collecting*. Slowly but steadily, some humanists appear to internalize this view as well, seeking refuge from the 'stamp collectors' label through quantitative and experimental methods. Digital humanities often serve as a bridge between disciplines, softening the rigid dichotomy between the sciences and the humanities – hence their occasional name, the *third culture*.[70]

The consequences of these changes might also include a reevaluation and shift in the hierarchy within the discipline itself, where advocates of new methods can build their discursive position by opposing what is perceived as traditional, implying that it is outdated, useless, and inadequate for today's challenges. One can also analyze this change in terms of the struggle for cultural capital in the academic field, as described by Pierre Bourdieu, who defines the clash between established figures and challengers in cultural production.[71] In this case, the contenders for a privileged discursive position in humanities are researchers questioning the traditional methods and confronting them with their own, more closely related to the exact sciences. The potential shift involves transforming the entire habitus of the researcher, also in the outsiders' eyes. On the level of stereotypical images and cultural representations, it includes replacing the figure of the great professor-sage, a polymath buried in books or wielding brilliant philosophical reflection, with the vision of a technocrat humanist – a rationalist, member of a research team, practitioner, and experimenter. Such a potential discursive shift will also have consequences for the material conditions of knowledge production; in the case of quantitative research in the cultural domain, those conditions often require expensive equipment and software or the creation of laboratories and well-organized research teams, traditionally the domain of the sciences. On the other hand, being less dependent on rhetorical strength and, consequently, less reliant on symbolic capital – including cultural and linguistic capital, which significantly determined hierarchies in traditional humanities – may partially level the playing field, slightly reducing the disparities between the peripheries and the discursive center of knowledge production and dissemination in humanities. Therefore, to understand the computational turn in contemporary humanities, it is insufficient to trace the latest methodologies, fresh discoveries, and leading trends. This will also require concepts and theories developed in the history of ideas, the sociology of knowledge, and critical discourse analysis.

Discursive embedding

We currently find ourselves in an intriguing phase of discourse – a moment where quantitative methods in humanities and film studies have already established certain traditions, academic credibility, and institutional grounding, yet their status and ultimate place in relation to traditional ways of producing and disseminating film knowledge remains uncertain. Especially digital humanities have developed a robust and stable infrastructure for knowledge production and dissemination – both tangible and symbolic – with roots dating back several dec-

ades. There are leading centers engaged in digital humanities, such as the Stanford Literary Lab at Stanford University, the Cultural Analytics Lab at the City University of New York, the Cologne Center for eHumanities at the University of Cologne, or metaLab – a joint initiative of Harvard University and Freie Universität Berlin – to name just a few, as well as organizations coordinating work in the field of digital humanities, such as the Association for Computers and the Humanities, the NEH Office for Digital Humanities, the Alliance of Digital Humanities Organizations, the European Association for Digital Humanities (EADH), or centerNet – an international network of digital humanities centers. Research results appear in dedicated journals, including *Computers and the Humanities Journal*, *Digital Scholarship in the Humanities*, *Journal of Cultural Analytics*, *Digital Humanities Quarterly*, *Big Data and Society*, or *Empirical Studies of the Arts*, and are presented and discussed at international conferences with significant traditions, such as the Annual Digital Humanities Conferences.

This field has also reached critical mass in the number of researchers and the advancement of their projects and methods, as evidenced by the emergence of various approaches, schools, and traditions, forming distinct methodologies depending on the preferences and backgrounds of researchers, as well as the objects of inquiry and the questions posed. The above overview of major trends in quantitative research in film studies remains brief and by no means exhaustive, but it reveals some of these approaches, such as cultural analytics or stylometric studies. One should also mention approaches situated within the broader framework of cultural evolution, which treats cultural transformations similarly to how biologists approach evolutionary processes in nature.[72] In other words, the field has become too vast for a single scholar to navigate freely, and the level of specialization in some research areas suggests significant progress taking place practically from year to year.

In recent years, popular knowledge based on this paradigm has also emerged. Alongside a multitude of specialized texts, research reports, and elaborate methods, sometimes assuming the form of lengthy monographs,[73] we have seen the first books based on reliable data yet with a clearly journalistic and popularizing character, including works like *You Are What You Watch* by Pulitzer Prize-winning author Walt Hickey.[74] Thus, we now have both advanced scientific research and accessible summaries for a wider audience. Still, we lack something in between – a good academic synthesis that not only acknowledges what we learn about cinema and its accompanying culture, production, or audiences through empirical and quantitative methods, but also provides meta-theoretical reflection. Its main aim should be the reintegration of quantitative approaches with traditional methods, thereby preventing a split within film studies between quantitative and qualitative researchers, embedding within the structures of knowledge production, and fostering appropriate critical and theoretical reflection.

All this should aim to propose a new, integrated yet multimodal model of quantitative methods and digital humanities in relation to the study of cinema, reflecting the changes in film studies and in cinema itself – increasingly more digitalized and informed by new technologies for production, distribution, and reception, including generative AI, streaming, and social media. Digital cinema

requires digital methods of research. As Giorgos Dimitriados puts it, *film production is perpetually on the verge of something new, a condition which more often than not produces corresponding dynamics for the research approaches that serve it*.[75]

It seems that the lack of systematic identification and interpretation of these dynamics might form the greatest gap at the moment. The discipline's further progress requires a two-pronged approach: on the one hand, organizing the developing methodologies and tendencies in quantitative film research and outlining their genealogy, and on the other hand, theoretical and methodological reflection on their possibilities, applications, and limitations. It is necessary to critically examine this entire turn, to embed it in the appropriate historical context of the discipline itself and, more broadly, in the transformations happening within the academia. In other words, we need to take a distant view at the practice of distant viewing.

---

[1] D. Berry, "The Computational Turn: Thinking About the Digital Humanities", *Culture Machine* 2011, vol. 12.

[2] G. Hall, "Toward a Postdigital Humanities: Cultural Analytics and the Computational Turn to Data-driven Scholarship", *American Literature* 2013, vol. 85, no. 4, pp. 781-809; B. Bozic et al., *Computational History and Data-Driven Humanities*, Springer, Cham 2016.

[3] D. Berry, op. cit., p. 2.

[4] M. Jockers, *Macroanalysis: Digital Methods and Literary History*, University of Illinois Press, Champaign 2013.

[5] F. Moretti, *Distant Reading*, Verso, London – New York 2013.

[6] L. Manovich, *Cultural Analytics*, MIT Press, Cambridge 2020, p. 9.

[7] C. Roth, "Digital, Digitized, and Numerical Humanities", *Digital Scholarship in the Humanities* 2019, vol. 34, no. 3, pp. 616-632.

[8] *The Datafied Society Studying Culture through Data*, eds. M. T. Schaffer, K. van Es, Amsterdam University Press, Amsterdam 2017; see also *Information and Culture* 2022, vol. 57, no. 1 (special issue on datafication and cultural heritage).

[9] E. Altenloh, *Zur Soziologie des Kino. Die Kino-Unternehmung und die sozialen Schichten ihrer Besucher*, Eugen Diederichs, Jena 1914.

[10] O. Bell, "Sociological Aspects of the Cinema", *Nature* 1939, vol. 144.

[11] S. Rowson, "A Statistical Survey of the Cinema Industry in Great Britain in 1934", *Journal of the Royal Statistical Society* 1936, vol. 99, no. 1, pp. 67-129.

[12] J. Spraos, *Decline of the Cinema: An Economist's Report*, George Allen and Unwin Ltd, London – Crows Nest 1962 (reissue: Routledge, London 2014).

[13] C. Arrowsmith, D. Verhoeven, A. Davidson, "Exhibiting the Exhibitors: Spatial Visualization for Heterogeneous Cinema Venue Data", *The Cartographic Journal* 2014, vol. 51, no. 4, pp. 301-312; D. Verhoeven, A. Davidson, B. Coate, "Australian Films at Large: Expanding the Evidence about Australian Cinema Performance", *Studies in Australasian Cinema* 2015, vol. 9, no. 1, pp. 7-20.

[14] C. Arrowsmith, D. Verhoeven, A. Davidson, B. Coate, "Kinomatics: A Global Study into Cinema Data", in: *Geospatial Science Research 3*, RMIT University, Melbourne 2014; B. Coate, D. Verhoeven, A. Davidson, "The Cinema Cities Index: Comparing Urban Cinema Cultures around the World", *Media International Australia* 2017, vol. 163, no. 1, pp. 163-175.

[15] See S. Loist, Z. Samoilova, "How to Capture the Festival Network: Reflections on the Film Circulation Datasets", *Necsus: European Journal of Media Studies* 2023, vol. 12, no. 1, pp. 363-390.

[16] B. Bohnenkamp, A.-K. Knapp, T. Hennig-Thurau, R. Schauerte, "When Does It Make Sense to Do It Again? An Empirical Investigation of Contingency Factors of Movie Remakes", *Journal of Cultural Economics* 2015, vol. 39, no. 1, pp. 15-41.

[17] J. Sedgwick, "From POPSTAT to RelPOP: A Methodological Journey in Investigating Comparative Film Popularity", *Journal for Media History* 2020, vol. 23, no. 1-2.

[18] P. Tinits, O. Sobchuk, "Open-ended Cumulative Cultural Evolution of Hollywood Film Crews", *Evolutionary Human Sciences* 2020, vol. 2, pp. 1-16.

[19] M. Stelmach, A. Hołobut, J. Rybicki, "Quantifying the Remake: A Historical Survey", *Journal of Adaptation in Screen and Performance* 2022, vol. 5, no. 3, pp. 211-226.

[20] J. Garncarz, *Wechselnde Vorlieben. Über die Filmpräferenzen der Europäer. 1896-1939*, Stroemfeld, Frankfurt am Main – Basel 2015.

[21] S. Loist et al., "Re-Framing the Picture: An International Comparative Assessment of Gender Equity Policies in the Film Sector", 2024, https://opus4.kobv.de/opus4-filmuniversitaet/frontdoor/index/index/docId/390 (accessed: 9.08.2024).

[22] M. Hagener, D. Roig-Sanz, "Digital Film Historiography: Challenges of/and Interdisciplinarity", *Journal of Cultural Analytics* 2024, vol. 9, no. 4.

[23] D. Verhoeven, "New Cinema History and the Computational Turn", in: *Beyond Art, Beyond Humanities, Beyond Technology: A New Creativity, Proceedings of the World Congress of Communication and the Arts Conference*, University of Minho, Braga 2012.

[24] See e.g., C. D. Reinhard, C. J. Olson, *Making Sense of Cinema: Empirical Studies into Film Spectators and Spectatorship*, Bloomsbury Publishing, London 2016.

[25] F. Fearing, "Influence of the Movies on Attitudes and Behavior", *Annals of the American Academy of Political and Social Science* 1947, vol. 254, pp. 70-79.

[26] See M. L. Anderson, "Taking Liberties: The Payne Fund Studies and the Creation of the Media Expert", in: *Inventing Film Studies*, eds. L. Grieveson, H. Wasson, Duke University Press, Durham 2008, pp. 38-65.

[27] D. Matyas, "You Had My Curiosity, but Now You Have My Attention: Camera Movement and the Spatial Distribution of Audience Attention in Cinema", thesis prepared at Yale University Department of Cognitive Science, 2016.

[28] E. Dayan et al., "Motion Cues Modulate Responses to Emotion in Movies", *Scientific Reports* 2018, vol. 8, no. 1; G. Raz et al., "A Robust Neural Fingerprint of Cinematic Shot-Scale", *Projections* 2019, vol. 13, no. 3, pp. 23-52; M. V. di Oleggio Castello et al., "An fMRI Dataset in Response to »The Grand Budapest Hotel«, a Socially-rich, Naturalistic Movie", *Scientific Data* 2020, vol. 7.

[29] J. Williams et al., "Cinema Audiences Reproducibly Vary the Chemical Composition of Air During Films, by Broadcasting Scene Specific Emissions on Breath", *Scientific Reports* 2016, vol. 6.

[30] B. Nemeth et al., "Bloodcurdling Movies and Measures of Coagulation: Fear Factor Crossover Trial", *British Medical Journal* 2015, vol. 351.

[31] A. D'Aloia, "The Intangible Ground: A Neurophenomenology of the Film Experience", *Necsus: European Journal of Media Studies* 2012, no. 2, pp. 219-239.

[32] U. Hasson et al., "Neurocinematics: The Neuroscience of Film", *Projections* 2008, vol. 2, no. 1, pp. 1-26.

[33] *Psychocinematics: Exploring Cognition at the Movies*, ed. A. P. Shimamura, Oxford University Press, Oxford 2013.

[34] R. Eugeni, "Neurofilmology: Semiotics, Cognitivism, Audiovisual Experience", in: *Open Semiotics. Volume 3: Texts, Images, Arts*, ed. A. Biglar, L'Harmattan, Paris 2023, pp. 421–434.

[35] A. B. Kovacs, O. Papp-Zipernovszky, "Causal Understanding in Film Viewing: The Effects on Narrative Structures and Personality Traits", *Empirical Studies of the Arts* 2019, vol. 37, no. 1, pp. 3-31.

[36] *Seeing into Screens: Eye Tracking and the Moving Image*, eds. T. Dwyer et al., Bloomsbury Publishing, London 2017.

[37] See M. Mera, S. Stumpf, "Eye-Tracking Film Music", *Music and the Moving Image* 2014, vol. 7, no. 3, pp. 3-23.

[38] D. B. Jones, "Quantitative Analysis of Motion Picture Content", *The Public Opinion Quarterly* 1942, vol. 6, no. 3, pp. 411-428.; J. Doniol-Valcroze, "Déshabillage d'une petite bourgeoise sentimentale", *Cahiers du cinéma* 1954, no. 31, pp. 2-14.

[39] B. Salt, "Statistical Style Analysis of Motion Pictures", *Film Quarterly* 1974, vol. 28, no. 1, pp. 13-22.

[40] The bibliography and the collection of articles written within the project is available at: https://cinemetrics.uchicago.edu/articles (accessed: 9.08.2024).

[41] M. Eder, J. Rybicki, M. Kastamont, "Stylometry with R: A Package for Computational Text Analysis", *The R Journal* 2016, vol. 8, no. 1, pp. 107-121.

[42] B. Adams, C. Dorai, S. Venkatesh, "Towards Automatic Extraction of Expressive Elements from Motion Pictures: Tempo", *IEEE: Transactions on Multimedia* 2002, vol. 4, no. 4; K. Pustu-Iren et al., "Automated Visual Content Analysis for Film Studies: Current Status and Challenges", *Digital Humanities Quarterly* 2020, vol. 14, no. 4.

[43] T. Arnold, L. Tilton, "Distant Viewing: Analyzing Large Visual Corpora", *Digital Scholarship in the Humanities* 2019, vol. 34, issue supplement 1, p. 1.

[44] Idem, *Distant Viewing: Computational Exploration of Digital Images*, The MIT Press, Cambridge 2023.

[45] B. Flueckiger, G. Halter, "Methods and Advanced Tools for the Analysis of Film Colors", *Digital Humanities Quarterly* 2020, vol. 14, no. 4.

[46] See e.g., M. Savardi et al., "CineScale: A Dataset of Cinematic Shot Scale in Movies", *Data in Brief* 2021, vol. 36; M. Savardi et al., "CineScale2: A Dataset of Cinematic Camera Features in Movies", *Data in Brief* 2023, vol. 51.

[47] N. Redfern, "Computational Analysis of a Horror Film Trailer Soundtrack", *Journal of the Japanese Association for Digital Humanities* 2022, vol. 6, pp. 11-20.

[48] A. Hołobut, J. Rybicki, "The Stylometry of Film Dialogue: Pros and Pitfalls", *Digital Humanities Quarterly* 2020, vol. 14, no. 4.

[49] W. Buckland, *Who Wrote Citizen Kane? Statistical Analysis of Disputed Co-Authorship*, Springer, Cham 2023.

[50] D. Bordwell, *On the History of Film Style* (2nd ed.), Irvington Way Press, Madison 2018, p. 272.

[51] L. Manovich, op. cit., p. 9.

[52] F. Moretti, "The Slaughterhouse of Literature", *Modern Language Quarterly* 2000, vol. 61, no. 1, pp. 207-227.

[53] M. W. Macy, "An Emerging Trend: Is Big Data the End of Theory?", in: *Emerging Trends in the Social and Behavioral Sciences*, eds. R. Scott, S. Kosslyn, John Wiley and Sons, New York 2016; A. Masso et al., 'End of Theory' in the Era of Big Data: Methodological Practices and Challenges in Social Media Studies", *Acta Baltica Historiae et Philosophiae Scientiarum* 2020, vol. 8, no. 1, pp. 33-61.

[54] T. Davies, M. Frank, "»There's No Such Thing as Raw Data«: Exploring the Sociotechnical Life of a Government Dataset, *WebSci'13: Proceedings of the 5th Annual ACM Web Science Conference* 2013, pp. 75-78.

[55] *"Raw Data" Is an Oxymoron*, ed. L. Gitelman, MIT Press, Cambridge 2013.

[56] See e.g., S. U. Noble, *Algorithms Of Oppression: How Search Engines Reinforce Racism*, New York University Press, New York 2018.

[57] W. Brown, "Politicizing Eye-Tracking Studies of Film", *Refractory: A Journal of Entertainment Media* 2015, vol. 25.

[58] G. Fossati, *From Grain to Pixel: The Archival Life of Film in Transition*, Amsterdam University Press – Eye Film Institute Netherlands, Amsterdam 2018.

[59] B.-C. Han, *The Expulsion of the Other: Society, Perception and Communication Today*, Polity Press, Cambridge 2018, p. 4.

[60] C. O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*, Crown Publishers, New York 2016.

[61] T. Smith, "Watching You Watch 'There Will Be Blood', https://www.davidbordwell.net/blog/2011/02/14/watching-you-watch-there-will-be-blood/ (accessed: 9.08.2024).

[62] D. Bordwell, "Hands (and Faces) across the Table", https://www.davidbordwell.net/blog/2008/02/13/hands-and-faces-across-the-table/ (accessed: 9.08.2024).

[63] R. Peels, L. Bouter, "The Possibility and Desirability of Replication in the Humanities", *Palgrave Communication* 2018, vol. 4.

[64] D. Berry, op. cit., p. 2.

[65] Ibidem, p. 9.

[66] A. W. Crosby, *The Measure of Reality: Quantification and Western Society. 1250-1600*, Cambridge University Press, Cambridge 1997.

[67] J. Gray, 'Let Us Calculate!': Leibniz, Llull, and the Computational Imagination", *The Public Domain Review*, https://publicdomainreview.org/essay/let-us-calculate-leibniz-llull-and-the-computational-imagination/ (accessed: 10.08.2024).

[68] D. B. Jones, op. cit., p. 411.

[69] C. P. Snow, *The Two Cultures and the Scientific Revolution*, Cambridge University Press, Cambridge 1961.

[70] A. Hunter, "Digital Humanities as Third Culture", *MedieKultur* 2014, vol. 57, pp. 18-33.

[71] P. Bourdieu, "The Field of Cultural Production, or: The Economic World Reversed", *Poetics* 1983, vol. 12, no. 4-5, pp. 311-356.

[72] A. Acerbi, *Cultural Evolution In The Digital Age*, Oxford University Press, Oxford 2020; in the field of film studies, see e.g., S. Sreenivasan, "Quantitative Analysis of the Evolution of Novelty in Cinema through Crowdsourced Keywords", *Scientific Reports* 2013, vol. 3; T. Grodal, "How Film Genres are a Product of Biology, Evolution and Culture: An Embodied Approach", *Palgrave Communications* 2017, vol. 3.

[73] A. Heftberger, *Digital Humanities and Film Studies: Visualising Dziga Vertov's Work*, Springer, Cham 2018; J. Cutting, *Movies on Our Minds: The Evolution of Cinematic Engagement*, Oxford University Press, Oxford 2021.

[74] W. Hickey, *You Are What You Watch: How Movies and TV Affect Everything*, Workman Publishing, New York 2023.

[75] G. Dimitriados, "Considering Eye-Tracking as a Validation Tool in Cinema Research", *Acta Universitatis Sapientiae, Film and Media Studies* 2021, vol. 20, p. 129.

Miłosz Stelmach Research and Teaching Assistant in film studies at the Jagiellonian University in Kraków. Editor-in-chief of the film magazine *Ekrany.* His main research interests are cinematic modernism, cinema of Eastern Europe, and transformations of contemporary arthouse cinema.

Bibliography

Berry, D. (2011). The Computational Turn: Thinking about the Digital Humanities. *Culture Machine*, 12, pp. 1–22.

Bordwell, D. (2018). *On the History of Film Style* (2nd ed.). Madison: Irvington Way Press.

Brown, W. (2015). Politicizing Eye-Tracking Studies of Film. *Refractory: A Journal of Entertainment Media*, 25.

Crosby, A. W. (1997). *The Measure of Reality: Quantification and Western Society. 1250–1600*. Cambridge: Cambridge University Press.

Davies, T., Frank, M. (2013). "There's No Such Thing as Raw Data": Exploring the Sociotechnical Life of a Government Dataset. In: *WebSci'13: Proceedings of the 5th Annual ACM Web Science Conference* (pp. 75–78). New York: Association for Computing Machinery.

Dimitriados, G. (2021). Considering Eye-Tracking as a Validation Tool in Cinema Research. *Acta Universitatis Sapientiae, Film and Media Studies*, 20, pp.129–150.

Gitelman, L. (ed.). (2013). *"Raw data" Is an Oxymoron*. Cambridge: MIT Press.

Gray, J. (2016). *"Let Us Calculate!": Leibniz, Llull, and the Computational Imagination*. PublicDomainReview.org. https://publicdomainreview.org/essay/let-us-calculate-leibniz-llull-and-the-computational-imagination/

Han, B. (2018). *The Expulsion of the Other: Society, Perception and Communication Today*. Cambridge: Polity Press.

Hunter, A. (2014). Digital Humanities as Third Culture. *MedieKultur*, 57, pp. 18–33.

Jockers, M. (2013). *Macroanalysis: Digital Methods and Literary History*. Champaign: University of Illinois Press.

Jones, D. B. (1942). Quantitative Analysis of Motion Picture Content. *The Public Opinion Quarterly*, 6 (3), pp. 411–428.

Manovich, L. (2020). *Cultural Analytics*. Cambridge: MIT Press.

Moretti, F. (2013). *Distant Reading*. London – New York: Verso.

Roth, C. (2019). Digital, Digitized, and Numerical Humanities. *Digital Scholarship in the Humanities*, 34 (3), pp. 616–632.

Snow, C. P. (1961). *The Two Cultures and the Scientific Revolution*. Cambridge: Cambridge University Press.

Słowa kluczowe:
metody ilościowe;
humanistyka cyfrowa;
zwrot obliczeniowy;
filmoznawstwo;
humanistyka oparta na danych

Abstrakt

Miłosz Stelmach

Kino się liczy. Zwrot obliczeniowy i metody ilościowe w badaniach nad filmem

Autor tekstu proponuje krytyczną analizę obecnych trendów i potencjalnych zastosowań metod ilościowych w badaniach filmoznawczych. W tym celu przeprowadza zwięzłą rekonstrukcję podstaw metodologicznych, ścieżek rozwoju historycznego oraz kluczowych osiągnięć narzędzi statystycznych, eksperymentalnych oraz tych z zakresu humanistyki cyfrowej w odniesieniu do badań mediów audiowizualnych. Obejmuje to zarówno przegląd rozwijających się obecnie zjawisk, jak i filozoficzny namysł nad źródłami, konsekwencjami i potencjalnymi ograniczeniami badań ilościowych w obszarach tradycyjnie zajmowanych przez nauki humanistyczne. Metody ilościowe nie są przy tym traktowane jako zamiennik istniejących paradygmatów, lecz jako ich uzupełnienie, rozszerzenie, a często także inspiracja. Pozwala to zrozumieć dokonujące się obecnie przemiany całej dyscypliny, ale również zintegrować nowe narzędzia z tradycyjnymi podejściami badawczymi oraz zidentyfikować pułapki i trudności związane ze zmianą paradygmatu.