Journal of Cultural Economics (2021) 45:171–211
https://doi.org/10.1007/s10824-019-09372-1

ORIGINAL ARTICLE

Leveraging analytics to produce compelling and profitable film content

Ronny Behrens[1] · Natasha Zhang Foutz[2] · Michael Franklin[3] · Jannis Funk[4] · Fernanda Gutierrez-Navratil[5] · Julian Hofmann[6] · Ulrike Leibfried[7]

Received: 3 February 2019 / Accepted: 11 December 2019 / Published online: 14 January 2020

Abstract
Producing compelling film content profitably is a top priority to the long-term prosperity of the film industry. Advances in digital technologies, increasing availabilities of granular big data, rapid diffusion of analytic techniques, and intensified competition from user-generated content and original content produced by subscription video on demand platforms have created unparalleled needs and opportunities for film producers to leverage analytics in content production. Built upon the theories of value creation and film production, this article proposes a conceptual framework of key analytic techniques that film producers may engage throughout the production process, such as script analytics, talent analytics, and audience analytics. The article further synthesizes the state-of-the-art research on and applications of these analytics, discuss the prospect of leveraging analytics in film production, and suggest fruitful avenues for future research with important managerial implications.

Keywords Entertainment analytics · Big data · Content production · Film producer · Film industry

1 Introduction

Producing compelling and profitable content is key to the long-term prosperity of the film industry. Content proliferation and intensified competition for viewers from Subscription Video on Demand (SVOD) platforms, whose success has often been attributed to leveraging analytics, have further escalated the desire for film producers

The authors contributed equally and were listed alphabetically. The authors thank the co-editors of this Special Issue, Thorsten Hennig-Thurau, S. Abraham (Avri) Ravid, and Olav Sorenson, and all other participants of the 2018 Mallen 20 Conference for their valuable comments. The authors also thank Josh Eliashberg, Sam Hui, and Ann-Kristin Kupfer for their valuable input.

✉ Ronny Behrens
ronny.behrens@wiwi.uni-muenster.de

Extended author information available on the last page of the article

to turn to analytics. While the role is difficult to narrowly define (Noam 2018), film producers oversee numerous aspects throughout the entire production phase, such as selecting scripts and talent, securing financing, and coordinating directing and editing. Essentially, they are the strategic enablers of creative ideas and organizers of the ideas' realizations (Lee et al. 2018). While independent producers may face distinct challenges compared to studio producers, the goal of producing compelling content that accomplishes financial success remains universal.

Digital technologies, such as digital filming, digital distribution, 3D production, augmented and virtual realities (AR, VR), have profoundly impacted the film industry and its broader competitive landscape. Examples of such impacts include content proliferation driven by SVODs' multi-billion-dollar investment in content production, reduced entry barrier to production, and user-generated content (Fontaine et al. 2018).[1] To compete in an environment where SVODs' revenues will rise to overtake theatrical revenues in 2019,[2] producers are increasingly turning to analytics as a promising route to produce compelling and profitable content (Bruneel et al. 2018; Fuselier 2017).[3]

Meanwhile, the substantial growth of digital consumption has generated a vast volume of granular individual-level data, propelling the diffusion of analytics across entertainment industries that conventionally operate upon creative intuition (Hennig-Thurau and Houston 2018; Fontaine et al. 2018). Across industries, such as SVOD, music, and gaming, the potential of analytics is well recognized and clearly demonstrated (Manovich 2018). For instance, Netflix has constructed one of the most sophisticated, analytics-driven recommendation engines and succeeded in micro-targeting the global audience (Gomez-Uribe and Hunt 2016; Kosterich 2016).

The academic literature has thus far concentrated on the downstream activities (i.e., the distribution sector) of the film value chain (Bloore 2009), often using the regression methods.[4] We instead focus on the upstream sector of film production from a film producer's perspective and those more advanced analytics, such as machine learning. Despite the focus on film production, the manuscript further incorporates the successful lessons from content productions in related audiovisual industries. This research also contributes to the literature by delineating a conceptual framework where analytics may be leveraged during film production, synthesizing the state-of-the-art literatures and suggesting directions for future research.

[1] Parrot Analytics (2018). Global digital original SVOD production trends. https://www.parrotanalytics.com/insights/global-svod-digital-originals-production-trends/.

PwC (2017). Consumer intelligence series: I stream, you stream. PwC. https://www.pwc.com/us/en/advisory-services/publications/consumer-intelligence-series/i-stream-you-stream/pwc-videoquake-i-stream-you-stream.pdf.

[2] McDonald (2018). Ampere: SVOD to overtake box office revenues next year. Digital TV.com. https://www.digitaltveurope.com/2018/12/17/ampere-svod-to-overtake-box-office-revenues-next-year.

[3] Fuselier (2017). Use these data-transparent companies to amplify your film distribution strategy. Sundance Institute Creative Distribution Initiative. https://www.sundance.org/blogs/creative-distribution-initiative/use-these-data-transparent-companies-to-amplify-your-film-distribution-strategy..

[4] Bloore (2009). Re-defining the independent film value chain. UK Film Council. https://www.bfi.org.uk/sites/bfi.org.uk/files/downloads/redefining-the-independent-film-value-chain.pdf.

In this process, we also address managerial responses and organizational arrangement relevant to film economics (Acheson and Maule 1994) by examining how the fine-grained, demand-side data can be leveraged by the supply side (Peukert 2018). By considering data as an emerging capability and toolset potentially aligning interests across development, production, distribution, and exhibition, the impacts of leveraging analytics on trade flows and vertical relations in the film value chain are explored (Chisholm et al. 2015; Aguiar and Waldfogel 2018).

Notwithstanding the potential offered by the richer data and advanced analytics, many challenges remain, stymieing their broader and deeper adoption by the film industry. For instance, most audience data are not readily or equitably accessible to producers. Another clear gap exists between academic research and industrial applications. Thus, this research takes an initial step to bridging this knowledge divide by surveying both the academic and industrial literatures and calling for more academia–industry collaborations.

The remainder of the article unfolds as follows: We first propose a conceptual framework of which and how key analytic techniques may be leveraged at each stage of film production. We then discuss the recent development of each technique before outlining the potential impediments to broader implementations of analytics in film productions and the promising avenues for future research. Throughout the manuscript, we do not advocate replacing creativity with analytics; instead, we promote their seamless integration. We also follow a probabilistic perspective, in which incorporating analytics does not deterministically guarantee a successful outcome, but rather increases the odds of success (Hennig-Thurau and Houston 2018; Foutz 2017).

2 Conceptual framework

Despite the diverse scholarly and industrial definitions, analytics commonly refer to the discovery, interpretation, and communication of meaningful patterns in data toward effective decision-making. These processes have found favor in the entertainment industries. Executives, such as the President of Warner Brothers' Worldwide Home Entertainment and Games, have highlighted the data's influences on film greenlight, promotion, and pricing.[5] Specific analytic techniques range from demand predictive analytics, social media analytics, to machine learning and more broadly artificial intelligence (AI) techniques, such as neural networks. Analytics in the film industry have largely found applications in the downstream value chain, such as ad planning and audience word-of-mouth (WOM) analysis (Hennig-Thurau et al. 2015). More recently, various SVOD platforms have leveraged analytics across much broader industrial processes, most controversially the creation of content itself. Nonetheless, producers' utilization of analytics remains an emerging area of research and practice.

[5] Prange (2018). Big data revolution. JCH Media Inc. https://www.mediaplaynews.com/big-data-revolution/.

In a hyper-competitive digital economy, persuading potential audiences to risk time and money requires "early" actionable insight from the perspective of producers and financiers. "Predictions made right before or after the official release … are too late for investors to make any meaningful decision" (Lash and Zhao 2016). Thus, an examination of how analytics may help in the "early" stages of content production will make a valuable contribution to the literature. To accomplish this, we propose a framework of leveraging analytics in film production by building upon the theories of value creation (e.g., Teece et al. 1997) and film production.[6] While the insights into film production guide us to dissect the production process into three primary stages, the value creation theory enables us to assess which analytics at each stage generate value and competitive advantages above the traditional approaches (Barney 1997). Specifically, drawing on the theories of film production, we divide the film production process into three major stages: development and packaging, production, and exploitation (Fig. 1). The first stage encompasses all decisions and actions leading up to a film's greenlight decision and before the cost becomes sunk in physical production. The second stage, production, involves everything that leads up to the final unalterable form of the film. And the third stage focuses on post-release marketing and learning processes that might impact future productions.

Although these stages, to a certain extent, present an artificial construct, they permit discrete decision points to be highlighted and relevant techniques evaluated. However, we should retain the consideration that there can be substantial leakage across stages, both of data pertinence and of decision practice. For instance, many forward-looking actions, such as investment, are based on past audience data, thus calling for cross-stage and cross-film analytics. Another case in point is ongoing content production, such as film franchises or TV seasons with extended life cycles and continual opportunities for analytics. An additional example is that granular audience data have been used to test a new cast member or assess who may boost audience ratings (Marolda and Krigsman 2018). Such adaptations of creative products based on analytics are not possible postproduction for films, despite that the marketing literature has clearly illustrated how distributors' targeting can be iteratively informed (Hennig-Thurau and Houston 2018; Foutz 2017). The implications for producers can come in amending the content emphases and right-sizing the budget of franchise iterations, bringing the exhibition data back to the greenlight considerations. Also, the iteration on script development within the development and packaging stage can be informed by data-led techniques.

Each of the above three stages involves several key decisions, presents distinctive challenges, and can benefit from intelligent applications of appropriate analytics. Table 1 displays our framework of the state-of-the-art essential analytic techniques that may be leveraged by producers at each stage to facilitate the creation of

[6] Theories of film production may be viewed at Bugaj (2013). The Production Pipeline Series. Private Blog. http://www.bugaj.com/?category=pipeline. Also see Marolda and Krigsman (2018). "Moneyball" for movies: data and analytics at Legendary Entertainment. https://www.cxotalk.com/episode/moneyball-movies-data-analytics-legendary-entertainment.

compelling and profitable content. In the next few sections, we will further discuss each technique.

To focus our inquiry into the use of data analytics on those aspects with a high potential to aid producers in creating superior value and establishing competitive advantages, we draw upon the recent developments in the value creation theory that focus on dynamic and innovation-driven market environments, such as the media industries (Hennig-Thurau and Houston 2018, pp. 128f.). Here, companies need to develop and maintain specific capabilities in order to adapt to these challenging circumstances. In particular, data- and analytics-related capabilities have been prescribed a paramount role for achieving success in light of intense competition (e.g., Wamba et al. 2017; Sharma et al. 2010). In this context, Verhoef et al. (2016)'s Big Data Value Creation Framework highlights the importance of owning suitable data, establishing adequate structures and competencies, as well as applying appropriate analytic techniques in order to create value for consumers and companies alike. This is in line with Bozik and Dimkovski (2019), who propose four central capabilities that facilitate successfully obtaining, managing, and analyzing data in commercial settings: (i) identification and obtainment of data (acquisition), (ii) data analysis and sense-making (assimilation), (iii) integration into existing knowledge (transformation), and (iv) application to commercial ends (exploitation).

Combining these capabilities with the above three primary production stages, we develop a framework of how and when producers may leverage data analytics to create compelling and profitable content of superior value to consumers and production companies (Fig. 2). At the end of each subsequent chapter, this framework will also serve as a guiding foundation for providing managers and researchers with a concise summary of the important capabilities enhanced as a result of analytics.

Despite numerous possibilities of leveraging analytics at various stages of film production, we argue that these efforts should be directed toward those that offer opportunities to create superior value, through addition to and not replacement of, traditional approaches. Netflix is a prime industry example that illustrates how creating such superior value in content planning and production, by gathering the right kind of data and analyzing them with the appropriate methods, can lead to sustainable competitive advantages over those less data-savvy industry players (Aguiar and Waldfogel 2018; Barney 1997). It gains a competitive edge in attracting and retaining consumers by leveraging viewing data to identify content compelling to the audience, as well as optimizing the efficiency of production operations by identifying bottlenecks and synergies from historic production data. This data-informed approach also allows Netflix to take calculated risks, provide valuable insights to producers, attract high-profile talent and creatives, and enhance the quality of the content library.

In line with Ravid (2013), who acknowledges the difficulties in decision-making based on past data due to the highly uncertain nature of predictions, research respecting these conditions is able to make useful statements about important areas of revenue distribution. We thus frame our investigations regarding data analytics as a new knowledge resource within such limitations, illuminating this resource's optimal management and allocation (Miller and Shamsie 1996).

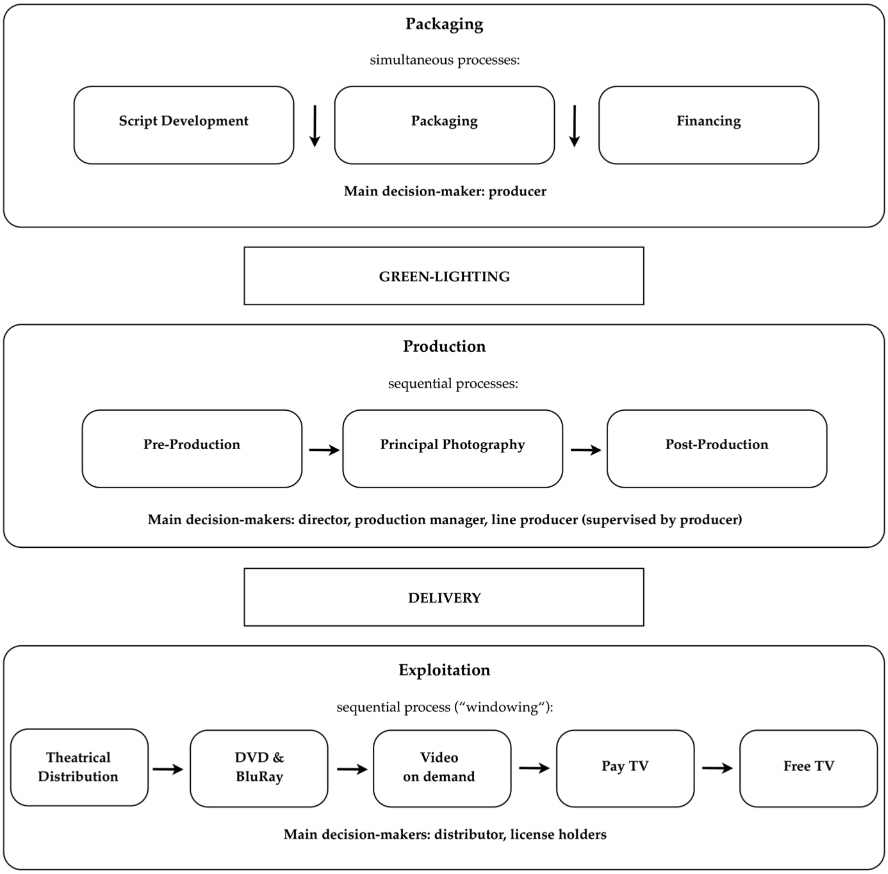

Fig. 1 Film production process

3 Development and packaging

Activities leading to the overarching greenlight decision that a film will be made often fall under the term "development and packaging." Such decisions typically rely upon many interrelated and iterative processes, including ideation, script development, casting, audience analysis, demand forecasting, and financing. Significant data-led innovations are impacting these tasks, but there is little academic literature on such data-led shaping processes, especially compared to the body of extant research on ex-post marketing (Simon and Schroeder 2019). Thus, important new toolsets will be examined in each subsequent subsection.

Table 1 Key decisions and analytics applications in feature film production

Key decisions and exemplary questions	Examples of present and potential applications of analytics
Development and packaging	
Ideation—what film to make?	Online viewing/search data mining for audience interest analysis
	Platform metrics for content development, e.g., non-viewing services such as reading platforms
Script—how to make the best film?	NLP to improve narrative structure, emotional tone—box office relations claimed with content characteristics
Casting—who will make the film compelling, who will sell the film to partners and audiences, which cast and crew will work best together?	Talent analytics
	Past performance (revenue/awards) and past collaboration for combination optimization
	Viewer data analyzed by cast variables
Demand forecasting leading to financing considerations combined with questions above leading to greenlight decision and structure—should we invest, and how much?	Financial structuring metrics and cost analytics—ROI complexity
	Demand prediction via micro-segmentation to the individual viewer
	Additions to toolset of regression methods and comparable title analysis
Future standardized data and use across partners for efficient decision-making; application of recommendation algorithms earlier into film life cycle; including individual consumer-level big data; application of granular data to producer-oriented research—including technical film finance questions of profitability	
Production	
Planning and coordination—location, schedule, budget, delivery—how can the film generate most ROI?	Operational production and postproduction process analytics (at early stage feed into financing)
	Where and how to organize a shoot—logistical optimization of talent availability, shooting time, and incentives
Postproduction analysis of consumer experience—what can be altered pre-release to optimize ROI?	Biometric viewer analytics
	WOM tracking over time
	Identification of emotional states and content highlights
	Edit film content to inform producer contribution to film positioning

Table 1 (continued)

Key decisions and exemplary questions	Examples of present and potential applications of analytics
Future dissemination of production analytics from market leaders as SaaS or platform to enable more uniform efficiency gains in pragmatic elements of production; assessment of viability and value of producer reaction to WOM pre-picture lock;	
Exploitation	
Re-evaluation to inform new productions—how do we do better next time?	Content analytics of franchise film reception to inform sequel development
	WOM and online search data analysis concerning film life cycle from festivals to awards and final revenues
Future WOM, online search data analysis concerning film life cycle—how do festivals, market sales, in-production marketing impact WOM, critical anticipation, awards selection, and performance?	

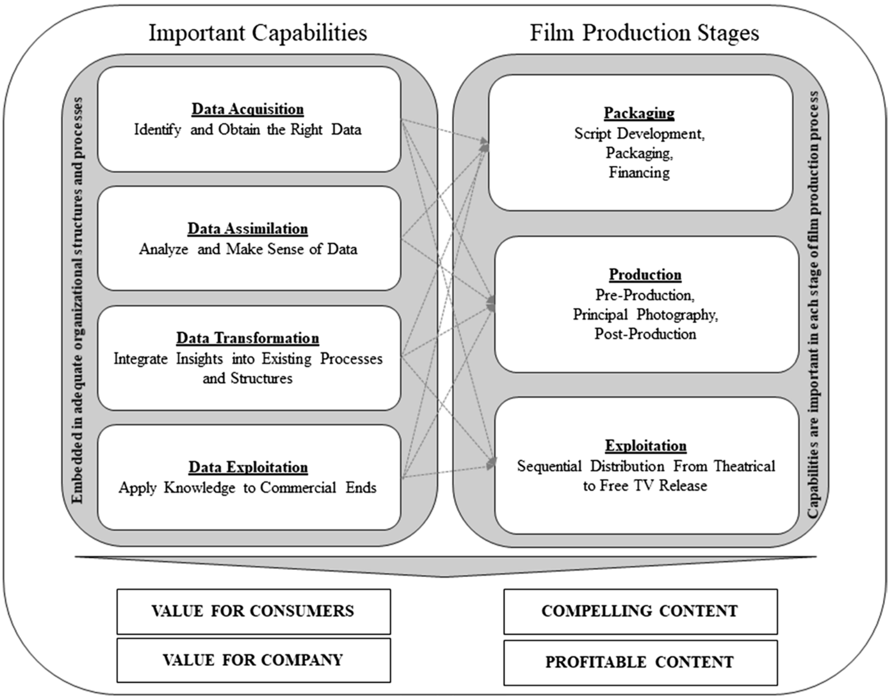

Fig. 2 Important capabilities during film production

3.1 Development: ideation analytics and script analytics

The challenge of determining a film's story most likely to be compelling to audiences and driving profits has led to the recognition of dominance and efficiency of sequels, remakes, and reproduced Intellectual Property (IP) (Ravid 2013, 2018). Recently, data-led approaches have opened up more opportunities for a producer in seeking and refining film ideas. Some capture and analyze data about a potential audience, so a producer may serve an uncovered demand, while others examine film scripts.

3.1.1 Ideation analytics

New tools and services for idea generation have emerged based on mining online content and its audiences. Broadly conceived, these tools aim to address the question of which kind of films to make, while also offer creative content analysis and script development capacities alongside ideation technologies. These innovation components can largely be separated into data-led tools that source content ideas (i.e., what story to tell) and revenue prediction tools based on script analysis (i.e., how to tell a story). Though these offerings can exist within one company, there are important distinctions to make.

Firstly, explicit claims of ideation via analytics are rare within industry. The use of analytics can even be explicitly disavowed by the most advanced practitioners of audience analysis. Technologies that challenge the primacy of traditional creative agencies are often feared and critiqued. As a result, their use can be hidden or framed as a decision support tool only. This can contrast with the possible use of some services, such as StoryFit that advertises its leverage of analyses of viewing histories, Web searches, and user engagement to identify the content that people actively seek. Using social media and broader consumer behavior data to infer audience interests in certain subjects and to generate content is also a component of forecasting, testing, and promotional efforts (Peukert 2018). Moving such practice upstream is important for producers' work and their understanding of the calculative process of potential partners.

Secondly, while producers for hire often deliver commissioned films for studios created "by the numbers" purely for profit, assessing what the audience want and catering to that demand have been largely based on the historical performance of the past films or related entertainment intellectual properties (IPs). And access to the wealth of online life and future-orients broadens the input parameters. This is especially important for addressing problems of historically underrepresented populations, where reliance on such past data does not change future content productions.[7] Wattpad, an online literature publishing platform, can be considered both as a more traditionally structured ideation service where popular writing can be discovered

[7] Smith et al. (2018). Inequality in 1100 Popular Films: Examining portrayals of gender, race/ethnicity, LGBT and disability from 2007 to 2017. USC Annenberg Inclusion Initiative. http://assets.uscannenberg.org/docs/inequality-in-1100-popular-films.pdf.

and then made into film and TV, but also an innovative one in terms of the scale, openness, diversity of content, and method of assessment used. The platform claims 65 million monthly users, 4 million writers, and 565 million stories, which are analyzed to determine why audiences think a story is important and then developed with TV networks or studios. Variables of read rates, read time, reader comments, search platform metrics, story metadata, genre, and storytelling mechanics are highlighted as the key analytic components to generate more compelling stories (Ramdarshan Bold 2016). However, the methods are proprietary and therefore opaque. Although built on collaborative filtering (CF) data science literature, the model is adapted to the contexts encompassing goals of completion and sharing—those engagement metrics deemed to allow for the observation of trends before the trends become obvious, rather than predicting tastes. The model has led to compelling hits for audiences in multiple media. Nonetheless, as we move chronologically down the value chain, from identifying a film's idea based on online activities or content popularity to developing that idea within a film framework, analytic details become clearer and richer. Scripts form the skeleton of the final creative product and as such are central to the industry's work and the application of analytics.

3.1.2 Script analytics

Two complementary analytic service elements focus on scripts for producers' development and packaging: the intended practice of improving a script itself, such as replacing clichés in a script with more nuanced ideas, and selecting which scripts to greenlight based on box office prediction. These goals are central to the services that apply natural language processing (NLP) and neural networks to script analysis (Napoli 2013).

Script analysis can function as a kind of film X-ray to identify elements, such as plot points or emotional tone, that drive success. Understanding the pathway to impact for such product characteristics can be important for producers. NLP has been used to find that the most talked about emotional arcs, not necessarily the most liked films, attain higher revenues.[8] That is, irrespective of budget and genre, the most financially successful films are those with an emotional fall preceding an emotional rise, and they are successful not because of that content characteristic solely and directly, but because they generate the most engagement. Thus, producing films most desired by the public, as measured by IMDB ratings, is not deemed a useful strategy; being liked is less useful than being talked about. While careful considerations of script, budget, genre, and tone are embedded in a producer's daily work with the goal of attracting positive attention, more specific content assessments potentially enable highly targeted producer action. Hunter et al. (2016) apply network text analysis to reveal that a screenplay's text network or concept map size strongly predicts its opening weekend box office. Eliashberg et al. (2007, 2014) further combine

[8] Del Vecchio et al. (2018). The data science of Hollywood: Using emotional arcs of movies to drive business model innovation in entertainment industries. Cornell University E-Prints Computer Science Computation and Language Resource document. https://arxiv.org/pdf/1807.02221.pdf.

NLP with domain knowledge when analyzing scripts to predict box office and guide greenlight decisions. Similar to Hunter et al. (2016), Eliashberg et al. (2014) find positive, significant predictors in the form of early exposition and a strong enemy. It is the use of data to describe a piece of work during creation that is relatively rare and sets such processes apart from the regular use of big data concerning the already produced work. While such indicators can add contextual details to producers' notes for writers, these analytics can also be helpful in relation to producers' use of extant greenlight support heuristics, for instance comparable titles analysis ("comparables"), and producers' understanding of their counterparties' operations. We argue that although the showcased approaches incorporating NLP still lack the semantic understanding of human experts, they can be a valuable asset for producers and complement producers' vast knowledge base by providing structural insights and uncovering success relations on a broader scale. With the recent, rapid advancements in NLP techniques, which aim at understanding semantic relations in textual data, how financiers or studios will evaluate a proposed script in the future presents a fruitful avenue for further research.

Producers can positively integrate script analytics into the packaging processes. Comparable titles analysis is usually used in greenlight decisions to provide supporting guidance regarding a film's likely performance. In order to help decision-makers determine whether to make a particular film, the process essentially asks "can a film like this be profitable?" Research shows that selections of comparables can be made more rigorous by including films with similar-sized main text network components, providing a further layer of reflection on likely revenues and informing casting and budgeting decisions accordingly (Hunter et al. 2016). Eliashberg et al. (2014) provide particular insights into testing their approach at a portfolio selection level, demonstrating that the economic significance (ROI) of their model's script selection is preferential to traditional, comparable titles analysis. Their script analysis makes a valuable addition to the assessment of genre, content variables, and semantic variables through the bag-of-words subset of textual variables, an NLP feature that can assess the overall tone of a film through key term frequencies. When considering the application of such data-intensive approaches from a producer's perspective, we must bear in mind both that such approaches can not only inform producers' review and guidance toward a screenwriters' work, but also influence how a potential investor may evaluate a film.

Challenges and future research opportunities regarding the development and packaging stage abound. Applications and research regarding ideation and script analytics face several challenges. When discussing the practical impact of their script analysis work, Eliashberg et al. (2014) note the challenges of data access, updating and scaling databases, and communication with studios. The issue with applications becomes even starker as methodologies developed by academics are leveraged primarily by major studios.[9] As a result, producers not working with

[9] Practitioners' use of academic methodologies is exemplified by Miguel Campo and his team at 20th Century Fox, who deploy methods from computer science, such as interactive intelligent systems, neural networks, and pattern recognition, to predict movie audiences using trailers. See Hsieh et al. (2018). Convolutional Collaborative Filter Network for video-based recommendation systems. Cornell University Computer Science Computation and Language Resource document. https://arxiv.org/

Motion Picture Association of America (MPAA) companies are unlikely to use these methodologies. Limited access to analytic techniques and their contextualized interpretations by the broader producer community restricts the flow of understanding across the value chain and aggregate efficiency. It is also vital to highlight that the use of such tools is only one of many considerations in group decision-making, and thus the empirical impacts on the industry are difficult to evaluate. Also, producers who partner with a studio are distinct from the continually employed executives at the multi-national companies who provide production financing and physical production services (Brookey and Zhang 2018; Noam 2018). The latter sits apart from the production, largely a temporary alliance of freelancers to ideate and realize the film, as opposed to an internalization of a whole value chain's activities within a company (Tashman et al. 2019). Even within integrated companies, such as major studios, data can be siloed.

Firm- and project-organization is highly important in the data analytics application context, and financing and arranging integrated production process have been highlighted as providing interesting puzzles for economic studies (McKenzie 2012). The divisionalization of filmmaking has been addressed as Corts (2001) notes that films with the same producer but different distributors do not benefit from efficient outcomes and that divisionalized companies tend to behave like integrated companies. Nonetheless, the role of modern analytics as a through-line in this field requires serious attention. Areas for examination include the role of data as a common language, the barriers to its use in different settings, such as contractual access or cross-division use, the existence or lack of collaboration across divisions and independent partners. These questions can provide further research streams adding to the literature's traditional attention to the US distribution sector and its connection with the marketing discipline (McKenzie 2012). Propagating the uptake and understanding of tools to enable engagement of analytics from multiple, integrated viewpoints, could support studios and wider approximation of integrated, cross-value chain datasets that trail the successful practices in China, where financiers, producers, ticketing apps and Internet platforms can be cross-owned to enable both marketing efficiencies and higher-level strategic decision-making.[10]

Footnote 9 (continued)

abs/1810.08189. Another example is Navaratjhna, Carr, and Mandt at Disney, who alongside their colleagues at Simon Fraser University and California Institute of Technology, utilize matrix and tensor factorization methods to examine facial expression data to model movie audience reactions (Deng et al. 2017). A third example is Movio, a third-party analytics provider serving STX Entertainment and the major studios. Movio stipulated that its causal inference methods were validated by Professor D. Rubin at Harvard University.

[10] Such innovations are covered in the trade press, such as Roxborough (2016). Alibaba Pictures buys into 'Mermaid' Financier Hehe Pictures. Hollywood Reporter. http://www.hollywoodreporter.com/news/alibaba-pictures-buys-mermaid-financier-hehe-pictures-950659.

3.2 Packaging: talent analytics

Building the right package and attaching suitable talent to a project is a core responsibility of any producer. Casting decisions are among the most important ex ante decisions that impact a film's success (Hunter et al. 2016). In industrial practice, two notable platform solutions facilitate the selections and packaging of talent. One is Cinelytic, a film decision support service that utilizes data about directors, actors, writers, and producers, in combination with historical box office, budget, and social media metrics, to derive proprietary talent scores or their "bankability." The other is Variety's V-score that is built upon the search volumes for specific talent.

The academic research has analyzed a variety of factors to guide casting decisions, such as the talent's individual track records and their past collaborative networks (Hofmann et al. 2017; Packard et al. 2016). A variety of analytic techniques are also developed to evaluate talent and their relationship to content (e.g., Kupfer et al. 2018). By analyzing, among others, WOM from essential stakeholders, primarily consumers, this stream of research often seeks to shortcut, if not replace, crucial human knowledge about the quality of the key talent and their team compatibility (Von Rimscha 2009). Other studies have also assessed talent's outward appeal and internal compatibility to forecast their impact on a film's revenue to guide casting decisions (Ghiassi et al. 2015; Hadida 2009). Alternatively, viewer behaviors, such as searching, rewinding, fast-forwarding, have also been utilized to guide casting decisions (Carr 2013; Kim et al. 2019a, b).

Beyond track records and viewer behaviors, talent's network and team composition are also explored (Cattani and Ferriani 2008; Cattani et al. 2013; Packard et al. 2016). For example, Packard et al. (2016) analyze the talent's collaborative network and reveal the importance of different network positions for front-of- versus behind-the-scene talent. Specifically, they find that higher eigenvector centrality which measures the connection with others more central to the network is critical for that front-of-the-scene talent, such as actors and actresses, whereas higher betweenness centrality that measures the capability to bridge sub-communities is more essential for behind-the-scene talent, such as cinematographers or set designers. Other studies have applied the network graph theory to films based on actor-specific characteristics and visualized topical structures and relationships among actors and film genres (Brandes et al. 2006; Haughton et al. 2014; Ghiassi et al. 2015). Kim et al. (2019a, b) further demonstrate how typecasting decisions can be driven by similarities among different actors. Moreover, Carr (2013) points out that Netflix annotates its content with hundreds of tags, the so-called metadata descriptors, in order to highlight good talent.[11]

Overall, the extant research is primarily prescriptive in nature. A common denominator of managerial relevance is the importance of team effects among the cast. The leveraging impact of successful team composition based on talent characteristics and synergy effects, such as talent-film fit, is demonstrated by analyzing

[11] New York Times (2013). Giving viewers what they want. https://www.nytimes.com/2013/02/25/business/media/for-house-of-cards-using-big-data-to-guarantee-its-popularity.html.

historical box office success (Hennig-Thurau et al. 2013). For instance, Brandes et al. (2006) show that data labeled according to film characteristics, such as genres and success figures, may help match actors using network-based algorithms. Such clustering techniques enable studio managers to identify groups of fitting actors that serve as a basis for successful compilation of teams. In a nutshell, this area of research, initially rooted in human resources, has revealed that successful applications of talent analytics are becoming a new source of competitive advantage (Chamorro-Premuzic 2016; Davenport et al. 2010). It has also given rise to the concept of "star power," which is among the most widely considered antecedents of box office success (Delen et al. 2007; Elberse 2007; Hur et al. 2016; Kim et al. 2015; Liu et al. 2014; Marshall et al. 2013).

3.3 Financing: demand predictive analytics and audience analytics

Financing involves processes of evaluation and composition that corral elements to complete the budget, effectively concluding the development and packaging stage with a greenlight to enter production. All stakeholders' greenlight decisions are directly influenced by a film's expected revenues. The well-known producer's challenge of the uncertain demand (De Vany and Walls 1996) is often formulated as "at what budget is there a sufficient audience for this film to greenlight it?" Thus, it is of critical importance to be able to forecast a film's demand before moving into the specifics of audience micro-segmentation, targeting, and reach. Nonetheless, demand forecasting is one of the longest standing challenges, particularly given the cultural diversity of the global film market (Dawar and Parker 1994), heavy-skewed revenue distribution (De Vany and Walls 2004; Walls 2005a, b), and subtle effects of competition, timing, and seasonality (Cabral and Natividad 2016; Dalton and Leung 2017; Einav 2007).

3.3.1 Demand predictive analytics

To better forecast a film's demand, a rich literature has examined the revenue impact of product characteristics (talent, budget, genre), marketing efforts (distributional width, advertising budget), third parties (critical reviews, consumer WOM), and audience traits (Hennig-Thurau and Houston 2018; Hadida 2009; Eliashberg et al. 2006; Hennig-Thurau et al. 2006; Elberse and Eliashberg 2003; Zufryden 1996). Both feature- and diffusion-based predictive approaches have employed a variety of regression methods (Hennig-Thurau and Houston 2018), which assume that common rules apply to a wide range of films. To address this untenable assumption, interaction effects have been incorporated, or a sample of films is split by genre for genre-specific analyses (Hennig-Thurau and Wruck 2000). On the other extreme, practitioners have argued that every film is unique, and thus no rules apply at all (Goldman 1983). The most common practice of the film industry is actually to take the middle ground, enlisting comparable films by assuming that there are patterns to discover, but only among similar films. Such scenario-based analogizing has been shown to outperform regression-based forecasting of film revenues (Lovallo et al.

2012) and remains conceptually sound since consumers hold relatively stable latent tastes for films, presumably based on underlying gratification-seeking behavior (Palmgreen and Rayburn 1982). The increasing availability of data and applications of analytics that detect consumers' preferences, moods, and feelings have also enabled film producers to optimize cross-cultural preferences for either local or global content (Chisholm et al. 2015; Peukert 2018).

The demand prediction literature draws attention especially when new data sources, such as social media (Asur and Huberman 2010), or new methods, such as prediction markets (Pennock et al. 2001; Spann and Skiera 2003; Foutz and Jank 2010; McKenzie 2013) or neural networks (Rhee and Zulkernine 2016; Ghiassi et al. 2015; Lash and Zhao 2016), are enlisted. Some of the latter have recently produced impressive predictive results on historical samples (Ghiassi et al. 2015; Lash and Zhao 2016). Nonetheless, these methods' lack of interpretability has presented challenges when used to convince financiers. Therefore, while a number of services, such as Epagogix, ScriptBook, and VaultML, offer early revenue predictions via machine learning, comparables continue to prevail as the dominant method for demand forecasting and to serve as the foundation of machine learning-based predictions (e.g., Eliashberg et al. 2014). Bruneel et al. (2018) draw a "sobering lesson," based on the state-of-the-art data mining models, that "it remains difficult to actually predict box office revenues with decent accuracy because of the presence of very strong outliers in the dataset." The impact of those outliers reinforces the fundamental dynamics of revenue distributions and the impracticalities of substantial reliance on such forecasting for investment decisions.

In light of this, when considering the producer's view, granularity is vital. However, scientific attempts to take return-on-investment (ROI) as a dependent variable in regression analyses usually define profitability in broad terms of box office revenues and budgets (Ravid 1999; Lehmann and Weinberg 2000; Miller and Shamsie 2001; Hadida 2003; Hennig-Thurau et al. 2001; Hennig-Thurau 2004; Ravid and Basuroy 2004; Jansen 2005; Lash and Zhao 2016). This approach is academically engaging, but less suited for practical applications from a producer's perspective, because the financial structure of films is so varied, and the way to determine a producer's or studio's ROI is complex. For instance, in casting, the effect of stars on revenues has been examined extensively (cf. Sect. 3.2), yet such a decision also affects the budget and may even change the revenue sharing scheme. Thus, the ways a casting decision affects the ROI are always multifold and indirect. From a producer's view, it makes more sense to estimate these effects separately and then assess the combined impact in a film-specific model, including the recoupment structure, instead of estimating an overall effect based on less granular data.

More recently, alternatives to historical norms of focusing on revenue performance have also been proposed, such as efficiency (Hababou et al. 2016) and formulations of profitability through different ROI calculations for greenlight modeling, including talent networks, release timing, and advanced content-specific features such as weighted average genre expertise (Lash and Zhao 2016). Another much more common heuristic to estimate profitability is a box office at twice the production budget to achieve break-even (Galvão and Henriques 2018), albeit overlooking global and ancillary revenues and advertising spends. Overall, the lack of direct correlation between revenues and profits (Lash and Zhao 2016) indicates that a more granular revenue breakdown will be necessary to comprehend the underlying processes on a deeper level.

3.3.2 Audience analytics

The increasing availability of individual consumer-level big data is poised to reconcile the two classic forecasting methods—a psychological approach via understanding an individual's motivation to watch a film and an economic approach via modeling the aggregate effects of film traits on the overall demand (Eliashberg et al. 2014). For example, recommender algorithms have proven successful at modeling individuals' film preferences (Jannach et al. 2016). As a result, forecasting individual demand is adding substantial value to identifying comparables, and most importantly, informing production decisions (Barbosu 2017). A notable industrial example is Movio's use of cinema loyalty program data to model an individual's propensity to see a certain new film.[12] Other examples include Jinni and JustWatch that offer marketers demand-side platforms to reach individual consumers based on their film tastes and illustrate how data can be leveraged at the global studio level to predict demand.

In summary, third-party decision support tools, such as Cinelytic, are adding analytics to the traditional industry heuristics of scenario modeling of profit and loss (PnL) accounts or control sheets (Townsend 2018). Nonetheless, industrial adoption and academic literature on modern predictive and audience analytics remain limited, partly due to the earlier, potentially over-ambitious, claims of analytics (Young et al. 2008). To address producers' concerns, more bounded ambitions have thus been forwarded.[13] As Hollywood and Silicon Valley converge through recent mergers, one may expect increased applications of such models and emergent embracing of "consumer science" at large. Lastly, a robust literature on the sources, workings, and contingencies of film financing (Ravid 2018; Debande 2018) explains debt and equity participation, adding to extant work that has detailed producer and investor intent and operation (Phillips 2004; Grantham 2012). However, examinations of how film risk is variable and subject to structural considerations have yet to be combined with the applications of analytics, which have been largely geared toward studios' franchising and genre strategies instead (Brookey and Zhang 2018). We hence invite more attention to higher level strategies that elicit opportunities for applying analytics to aid financing decisions in general. Specifically, we identify a more in-depth examination of forecasting impacts on revenue flows and recoupment structures as a promising field for further research, as seamless integration of analytic tools with

[12] Smith and Beguely (2017). Movio view: propensity, machine learning, automation, and why it matters. Movio Blog. https://movio.co/blog/movio-view-propensity-machine-learning-automation-and-why-it-matters/.

[13] For example, StoryFit (2018) caveats its proposition to "using big data to improve greenlight" with the proviso that "it is not possible to 'predict success' in the way that most people mean."

industry practices will be key to developing early research findings into applicable tools that may be leveraged as an actual competitive advantage (Table 2).

4 Production

Production encompasses pre-production (from greenlight to the first day of shoot), production (principal photography), and postproduction (from the last day of shoot to delivery of the film for distribution).

4.1 Pre-production and Production: Planning and Coordination Analytics

Production planning and coordination are pragmatic aspects of filmmaking often underrepresented in the marketing and management literatures. Yet the integrated production processes are noted to provide engaging topics for economics research (McKenzie 2012). While for instance, cinematographers and visual effect (VFX) providers are hired by producers, the impact of analytics on those activities is mainly felt by those selected service providers themselves, as opposed to the producer. Therefore, we focus on the aspects that reflect a producer's core concerns. Taking the individual film as the organizing unit of the industry, the economic logic and methods of performance have been understood in terms of project management (Gaustad 2018). It is from this perspective that analytics-led changes to the producer's existence can be usefully viewed.

As highly definable and measurable aspects of filmmaking, planning and coordination afford immediate opportunities for leveraging analytics and big data to address profitability issues, such as in budgeting, territory, and set selection, photography schedule optimization, workflow planning, and human resource management. For instance, film budgeting that involves cost and logistics analysis typically takes place over multiple iterations among various potential partners before greenlight and is thus in some ways interwoven with previous subsection activities. As part of this, evaluating shooting locations is a critical example of analytics' application to production planning and coordination. While line producers and studios have great experience, factors including unique project requirements, emerging tax incentives, and changing cast and crew competencies make this a highly dynamic and complex data environment. Creative intuition, experience, and expertise continue to be employed. But as more defined inputs, such as subsidy cost calculation (Ravid 2018), become more optimizable due to analytics, improved performance is enabled. Major studios have entire departments dedicated to this work with private databases on incentives, studio rates, wages, and currency risk per country (Castendyk 2018) and therefore enjoy the advantage of scale. Nonetheless, they have not invested in analytics comparable to SVODs (Cohen 2017).

Attention to the example of Netflix provides insight on the industry's leading edge that illuminates challenges and opportunities for the broader production sector and attendant research. Netflix is highly developed at data-led production, not least due to the substantial scale resulting from an $8 billion per annum content spend.

Table 2 Important capabilities at the packaging stage

Data acquisition	In ideation and script development, structured story data (e.g., genre, story components and storytelling mechanics) can be very helpful in combination with consumer interest/engagement data (e.g., Web searches and online interactions on social media) and historic film performance data
	For finding the right talent during packaging, both data on the individual level (e.g., commercial performance in past movies and social media/Internet performance metrics) and on the group level (e.g., synergies between cast members) can offer valuable guidance
	For financing, data on the product (e.g., budget and genre), marketing strategy (e.g., screens and advertising) and third parties/target audience traits (e.g., critical and consumer WOM) can be utilized to forecast a movies performance—however, data availability and thus the corresponding approach critically hinge on forecasting timing
Data assimilation	In case of ideation and script development, mostly textual analysis and neural networks are applied in order to gain insights into unstructured textual data and to establish a link with future success metrics
	For finding the right talent that fits the movie, descriptive- (to identify star power) and regression-based (to explore fit with film at hand) method can be employed and supplemented by network-based approaches in order to identify optimal cast composition
	To forecast future performance of a film, both feature- and diffusion-based regression models and modern machine learning techniques can be utilized
Data transformation	The discussed data and methods will help to improve existing processes and supplement managerial intuition on a broad spectrum of tasks. This ranges from the initial conception of a film (e.g., identifying the right stories and telling them in an optimal fashion), choosing the right talent (e.g., individual cast members and cast composition) to predicting the commercial success of a movie, which ultimately ease greenlight and production (planning)
Data exploitation	Value for consumers: By expressing or signaling their interest in certain stories, more non-stereotypical and nuanced movies that deviate from the four-quadrant targeting norm might be produced, offering consumers more compelling content that they desire
	Value for producers: In this stage, data analytic capabilities will aid producers to sift through available story material and identify potential hits more easily, and also guide them in selecting synergetic talent who both resonates with the audience and fits the story at hand. Ultimately, data analytics will add empirical evidence to the mostly experience-driven greenlight and financial discourse, potentially leading to a higher probability of producing profitable content

Yet even in the era of a billion data rows per second (Manovich 2018), data sparsity is still a challenge due to film's complexity as a cultural product and the inherent scope for substantial diversity across parameters of genre, budget, tone, cast, and resulting labor and equipment. Planning options for new films rarely have much like-for-like past data, so multiple inputs must be collated and combined. Netflix uses hierarchical models to afford production executives what-if scenario analyses of different location options. Since entering production with self-managed originals, Netflix has reengineered scheduling processes that were only recently paper-based or conducted via unsophisticated digitized systems (Walraven 2018). Automated optimization models for shooting schedules form planning and budgeting guides by very quickly assessing combinations of days, scenes, locations, talent availability, and cost type. The aim is to optimize studio performance through actionable metrics for quick decisions and reduce operational burdens not replacing creative freedoms (Walraven 2018). Just as there is a creative evaluative element to the cost estimation problem—the shooting location must be able to serve the story—the expertise of the producer remains in scheduling, but significant legwork is already done, increasing efficiency.

In summary, examining the gaps in data and skill access between the leading companies and typical producers not only demonstrates disparities in resources applicable to individual projects, often orders of magnitude apart, but also highlights a related exclusion from the ecosystems of evaluation that structure the filmmaking economy. As Hennig-Thurau and Houston (2018) point out, an entertainment company's management success is not only a function of the company's own role, about which it can make strategic decisions, but also concerns the role of other companies. Abilities to profitably match creative artistry with business discipline can be inhibited when interrelations between companies are increasingly predicated on unevenly distributed resources, such as analytics capacity and scaled resources to leverage such.

In the meantime, the increasing digitization of project management support as software as service (SaaS) tools, for instance, Movie Magic Scheduling, Vista Accounting, and Digital Production Office for crew contracts and payroll, opens up some efficiency gains to most producers. However, it is in the proprietary systems for globally scaled operations that the greatest advances exist, dependent on financial and time resources to acquire data sources and iterate across projects and territories. This has knock on impacts for other producer types, particularly in their understanding of how potential partners operate. For instance, Netflix applies machine learning to asset review editing and delivery processes. To operate these systems, there are delivery obligations on suppliers, and the performance can be validated and given a metric.[14] The difference in openness and connectivity between Netflix and its partners in different contexts of the production sector supply chain is important when considering producers' use of analytics. Where companies provide a continuous or regular support service for a fee, knowledge sharing and integration is key and performance is publicized, for instance, as redelivery rate metrics. Where involvement is sporadic and on different terms than established industry practice, such as buyout instead of incremental royalties, data sharing and systems integration are greatly reduced. Although Netflix has indicated they will be revealing more granular reporting, the precise nature of such data and cross-calculations needed by a producer to understand the nature of a project's performance and to leverage future deal-making remain extremely unclear. It is thus an important and challenging area for future research when considering competitive advantage.

[14] See Mosisoglu (2017). Netflix content data: From script to screen. Netflix Data Presentation 8/6/2017. https://www.youtube.com/watch?v=qXo9jTxfqJ8&t=2s.

4.2 Postproduction: experience analytics

Prior to release, a pretest or test-screening among a small sample of viewers can gauge viewers' experience. Experience analytics in our context encompass processes and technologies that provide insights useful for the producer during the editing process of a film and design of promotional materials.

There are a variety of complementary methods used to determine the emotional state of viewers and to detect personal highlights of films. Besides traditional focus groups and surveys, biometrics is also employed to assess experiences. Specifically, real-time physiological data, such as heart rate, blood pressure, brain activity, body temperature and perspiration, can be captured using wearable sensors. Moreover, body movements, facial expressions, eye tracking, and pupillary responses can also be recorded during film viewing (Bilasco et al. 2015). These psychophysiological channels have been shown to relate to affective responses to visual content. For instance, Abadi et al. (2013) investigate the effectiveness of different psychophysiological affective signals, and combinations thereof, to detect viewers' engagement levels while watching a film. They find that galvanic skin response and electroencephalogram provide comparable contributions, while facial motion tracking provides complementary information; thus, the best result is achieved by using all three modalities together. Soleymani et al. (2012) present a user-independent emotion recognition method using electroencephalogram (i.e., record of brain activity), pupillary response, and gaze distance, which outperforms individual self-reports for arousal assessments without underperforming on valence assessments. Their emotion recognition method offers real-time evaluations instead of retrospective evaluations, hence capturing viewers' "real" instead of "stated" reactions, which may suffer from demand effect or inaccuracy. Other advantages of using biometrics include eliminating the need to ask viewers any questions which may interfere with the ongoing experience itself and eliminating the need to decompose a product (film here) into multiple attributes as in Conjoint Analysis. Then, biometrics provide a good alternative to traditional surveys where emotions are self-reported, or expressing emotions with words or remembering the different emotions at each moment of the content is challenging. During film editing and postproduction, producers may thus leverage the insights generated from these metrics to detect viewer interest or engagement at different points of a film. For instance, the producer may adapt parts of the content according to what the audiences expect or desire to watch, change the title before release, or even reshoot parts of the film. The digitized test-screening environment allows quicker and more analyzable feedback and follow-up capabilities for producers who can then refine their audiovisual content accordingly.[15]

[15] There are examples of established market leaders with in-house practices, such as Legendary Entertainment. There are also emerging research-led endeavors. For example, ETC's Theatrical Demo Data Project tests a new in-theater technological solution to gather large-scale viewing data in real time (cf. Bergquist 2017. How my team and I are trying to revolutionize Hollywood. Medium Corporation. https://medium.com/@punkstrategy/how-my-team-and-i-are-revolutionizing-hollywood-9930e28937d).

Experience analytics also play a crucial role in guiding the production of marketing materials, for instance, trailers for domestic and international audiences. Analytics enable producers to identify the funniest or the most thrilling parts of a film, and even automatically extract the content (Hanjalic and Li-Qun 2005). On a related note, new analytics-propelled optimization procedures have emerged to automatically shorten film trailers as exemplified by Liu et al. (2018a, b). In cooperation with Netflix, the authors utilize Web-based facial expressions to identify the trajectories and especially peaks of happiness according to Ekman and Friesen's (1978) Facial Action Coding System using comedy films. Joining those data with consumers' corresponding trailer evaluations and stated intention to watch the film, the authors utilize a Bayesian backward-propagation algorithm to shorten the trailer to its essential parts. Their results indicate that this approach seems to be an effective and scalable way to produce marketing materials in a specific context. Nonetheless, future research is needed to investigate the role of other emotions and robustness with other film genres.

Another aspect to be considered when editing the content or producing a trailer is to maintain the coherence of the plot. Since interpretations, analyses, and predictions are related to the order of the content, a coherent and interesting plot drives consistent reactions and makes individual brains behave more similarly. Barnett and Cerf (2017) measure the levels of neural similarity by collecting the neural data from the moviegoers in a commercial theater. The authors use a 32-channel portable electroencephalography (EEG) system to record the brain activities of the participants. Then, the cross-brain correlations (CBC) as moment-to-moment synchrony in the EEG data across the study's audiences experiencing the same trailer are calculated. Also, two surprise surveys are conducted to assess the participants' content recall immediately after viewing the trailer and 6 months later. The study shows that the movie trailers that drive similarity in neural processing among participants enjoy higher levels of recall and tickets sales. Regardless, more research is needed in order to understand which trailer characteristics and execution factors lead to such a similarity in neural processing in order to provide actionable insights to producers.

During the pre-release stage, Simon and Schroeder (2019) highlight the general use of big data to edit the final cut toward the audience's inferred demand. Such pre-release applications of analytics are aligned with our recommended research agenda, which may benefit from re-orienting specific marketing techniques to the production, particularly early production, processes. Thus, we argue that although the previously discussed studies focus on the reception of film trailers, their core aim—to optimize consumers' experiences by evoking the right emotions at the right time—is also applicable to the editing process of the complete film materials. The application of such techniques can supplement existing approaches like test screenings by offering more granular insights into consumers' film reception, thus increasing the viewing intentions and ticket sales.

Lastly, the extant literature has focused on individual consumption and preference, although film viewing is inherently a social experience (e.g., Ramanathan and McGill 2007). Also, viewers vary in their knowledge and commitment, making the group experience and post-experience evaluations even more complex (Foutz 2017).

We hence deem the social experience analytics a fruitful avenue for further research, which will help producers take such important dynamics into account (Table 3).

5 Exploitation: WOM analytics

At this post-release marketing and learning stage, to enhance intermediate strategic decision-making, producers should direct their attention to one of the valuable assets of WOM data. WOM analytics represent an important staple in both the marketing literature and successful practice of film release management (Hennig-Thurau et al. 2015). With an increasing number of widespread online WOM sources, such as personal blogs, review sites, and instant messaging sites, consumers are increasingly turning to these sources due to their information quality and source credibility (Yeap et al. 2014). Nonetheless, it is not a trivial task to make sense of the immense amount of structured data (e.g., star ratings) and unstructured data (e.g., textual reviews). Earlier research has focused on straightforward metrics, such as volume (i.e., amount of WOM) and valence (positive or negative sentiment). However, mixed findings prevail regarding their relationship with sales. While some studies point to a positive impact of WOM volume (Liu 2006; Duan et al. 2008), others suggest WOM valence as the driver of box office success when controlling for pre-release advertising (Chintagunta et al. 2010). As a result, the immediate pertinence of this research for producers remains less clear-cut, as it lacks granular details into which aspects of a movie have created those consumers' responses. Still, distributors are provided with insights into markets that are most responsive to advertising, blog volume, and valence, and thus should be targeted to improve the release strategies (e.g., Gopinath et al. 2013).

For this reason, as reflected in the literature, insights derived from WOM analytics are mainly applied by distributors in the design and management of marketing and release campaigns (Hennig-Thurau et al. 2015; Hennig-Thurau and Houston 2018). Decision-makers who are aware of the importance of consumer WOM after release start the effective management process early in the lifecycle of a film, i.e., prior to its production. Usually, the corresponding task is delegated to those departments with the skills, resources, ongoing responsibilities, and contractual and effective dominion. Lee et al. (2018) explore the manipulation of sentiment on Twitter, recognizing film marketers' incentives to influence online information prior to release accordingly. However, considering the specific decisions and contextual framing of a producer's work, there is less clarity and detail with regard to the insights to be gained and applications to be derived from the ubiquitous WOM data. An important distinction is to be made from a related, but preceding concept: pre-release consumer buzz (Houston et al. 2018). Deemed an important precursor for a film's initial success, it is created in response to studio and distributor actions (budget, age restriction, distribution strategy, star power, and related creative works) and inferred information. But in contrast to WOM, it is not built upon real consumption experiences. Although it loses power when post-release WOM takes over (Houston et al. 2018), it is of high importance in a film's (and especially a blockbuster's) front-loaded diffusion (Hennig-Thurau and Houston 2018, pp. 109–113). And

Table 3 Important capabilities at the production stage

Data acquisition	Data for process optimization can be collected on a company level, recording central variables in past and present projects, or be bought from third-party database providers or pooled with similar companies
	In postproduction, data on the individual level become more relevant, encompassing responses and biometrics from, e.g., interviews, surveys, and laboratory studies
Data assimilation	With the help of SaaS production logistics tools or self-deployed machine learning algorithms, which focus on optimizing certain processes based on the underlying data and suitable process performance metrics
	For quantitative and biometrical response data, a wide range of quantitative methods, from standard regression techniques to Bayesian back-propagation algorithms, can be applied, depending on desired outcome
Data transformation	Support producer expertise in highly dynamic areas of production such as territory selection or workflow planning, utilizing data to improve and optimize information typically used in decision-making
	In postproduction, data about consumers' reactions to different versions of a film or different promotional material will supplement creative intuition in making critical adjustments before release to tailor the film more closely to target groups' taste and expectations. Furthermore, following the production work, algorithms can help to automate certain promotional decisions on a scale, e.g., automatically shortening trailers based on peak emotional moments (Liu et al. 2018)
Data exploitation	Value for consumers: While consumers might not directly benefit from production optimizations, the inclusion of their feedback in postproduction ultimately leads to films that more directly tailored to their needs and expectations
	Value for producers: Optimizing the production pipeline itself offers many opportunities to create synergetic effects and save costs, which directly affects a film's bottom line. Incorporating consumers' reaction in postproduction adds guiding evidence to creative adjustments to the film and makes marketing efforts more effective (better marketability through consumer insights) and efficient (through automation)

analyzing the antecedents and drivers of pre-release buzz from historic examples will offer an attractive avenue for research with impactful practical insights for film producers.

More recent examples of film WOM analytics and their applications by major companies (Yang and Yecies 2016) pertain to content customization and promotion strategies based on the derived audience preferences and behavior predictions. That is, while WOM analytics for content production are generally conducted by major SVODs, such as Netflix, Sina Video, and Tencent Video, they are rarely detailed in the film industry (Yang and Yecies 2016; Anfer and Wamba 2019). Similarly, the academic literature on such techniques and potential avenues for value creation in the production process remains sparse. Future research in this area is needed to offer critical insights to practitioners and academics alike, increasing the insights that can be derived from such readily available data. The dire need for such insights becomes clear when considering an example that is vital to the current portfolio of producers: sequels. Here, a sequel shares and continues some elements of the previous franchise, such as settings, characters, and major plot developments. WOM and revenue figures of the previous film are of high strategic importance (Hennig-Thurau et al. 2009; Yeh 2013), providing valuable information about what consumers (dis-)liked and wish to see in the sequel(s). Utilizing appropriate analytic techniques, such as topic modeling (e.g., Humphreys and Wang 2017), can help producers extract such insights on a scale and better assess a sequel's commercial potential and improve ideation and packaging processes.

Driven by the increasing digitalization and platformization of the film consumption experience, other data sources, such as the meticulous information about viewer behaviors, have surfaced and may offer further insights into what makes a film profitable and compelling. However, these data are not always available to producers and scattered across different market players, making assessments of these data's potential value to producers difficult. Academic and industry coverage also remains scarce. Furthermore, producers now have to take into account how other market players, such as studios, financiers, SVOD services, sales agents, or international distributors, use such data to evaluate the derived demand (Finch et al. 2015) in the development and packaging stages in order to assess these data's value to future film productions throughout the exploitation stage in often opaque ways.

SVOD companies play a big role in this development of increasingly available, yet non-transparent, data at the exploitation stage. Their uses of recommendation systems are examined (Van Roy and Yan 2010), but the challenge is daunting for a producer to understand how to give his/her film the best opportunity to be profitable, often without access to such data and corresponding analytic capabilities (Hennig-Thurau and Houston 2018). Compared to a producer's twentieth-century concerns for value generation at the exploitation stage, such as ensuring cast interview availability and high-quality still images for delivery to a sales agent, the current black-boxed analytics of downstream value creation can seem to encompass an entirely different industry. For instance, Netflix leverages AVA, a set of tools and algorithms, to explore around 8500 frames of content per hour in order to connect content and the right target audience that finds it compelling with the help of A/B testing. To achieve this and make sense of this vast amount of unstructured film data, AVA automatically annotates and ranks frames on aesthetic, creative, and diversity objects, including such factors as actor's facial poses, sentiment, and camera motion.[16] Yet again, producers having no transparency on how film characteristics, such as genre, cast, style, and tone, impact these evaluations that are intertwined with re-licensing and future project involvement (Lobato 2018). As a result, producers suffer from substantial information asymmetries (Opitz and Hofmann 2014). In contrast to the previous economic primacy of traditional theatrical release windows, where consumer downside risk varies significantly between heavy and casual attendees and producers derived valuable insights from utility analysis and postmortems alone (Campo et al. 2018), these exploitation data become less impactful with the increasing importance of SVOD services. The requirements for successful data acquisition, assimilation, transformation, and exploitation change, with lower stakes for individual content

[16] Riley et al. (2018). AVA: The Art and Science of Image Discovery at Netflix. Technology Blog. https://medium.com/netflix-techblog/ava-the-art-and-science-of-image-discovery-at-netflix-a442f163af6.

evaluation and increased importance for service attraction when assessing which content is commissioned, financed, and acquired (Campo et al. 2018). Thus, investigating what constitutes value and which factors contribute to making content profitable and compelling along this journey is a key area for future research.

6 Discussions

Taking a producer's perspective, we survey the academic and industrial literatures on the key analytic techniques leveraged throughout the elaborate, multistage film production process. We aim to capture the present state-of-the-art and suggest promising avenues for future research. In the next few subsections, we offer further discussions on the prospects of leveraging analytics in film production for both academic research and managerial implications.

6.1 Film producers need to leverage analytics to compete more effectively in the global content market

Global entertainment and media revenues will continue to rise with a 4.4% AAGR over the next 5 years, reaching $2.4 trillion by 2022 (PWC 2018). To compete more effectively in this lucrative global marketplace, where SVODs' revenues will overtake box office revenues in 2019, film producers need to leverage analytics to create the content appealing to consumers and profitable to production companies. Our review of the industrial practice, including SVODs' successful engagement of analytics, points to the urgency for film producers to leverage analytics in the increasingly data-driven competitive environment. Our review of the academic literature also demonstrates the present and potential value of analytics throughout the film production process. In summary, this research suggests that, consistent with value creation literature (Bozik and Dimovski 2019; Verhoef et al. 2016; Teece et al. 1997), film producers and production companies need to develop and maintain a set of crucial capabilities in order to survive and thrive in the highly dynamic environment characterized by intense competition and constant need for innovation.

With the increasing digitalization of content production, provision and consumption, developing the right data analytic capabilities will help producers complement traditional industrial approaches to ultimately establish and sustain competitive advantages by creating superior value for both themselves and consumers. Yet, in order to translate this wealth of data into the production of more compelling and profitable content, it is paramount to develop such capabilities in a systematic fashion, spanning from acquiring the right data, utilizing suitable analytic techniques, integrating the new insights with the previously available information and established process, to applying the newly generated insights to the commercial ends. As we previously showcased (also cf. Tables 2, 3 and 4), when done rigorously and on eye level with traditional approaches, these capabilities hold great potential to enhance almost every step along the film production value chain. In particular, we wish to highlight the following four areas with great potential for future academic research and industrial applications, especially in light of the film studios' position in today's competitive landscape, such as their relative focus on franchises and tent-pole movies, and the unique experiences entailed by theatrical films relative to SVOD home entertainment.

6.1.1 Content analytics

While conventional analytics have focused on the structured data often flowing from the operational systems, newer analytics will integrate more unstructured data, such as text (scripts, screenplays, audience reviews), image (posters, ads, packaging), voice (recordings), and video (films, trailers) (Toubia et al. 2019; Liu et al. 2018a, b). Content analytics are particularly important for major studios as they tend to pursue a tent-pole strategy with high stakes. Content analytics will offer producers much more granular insights into the success potential of a script, trailer, or film, as well as real-time guidance on how to adapt the content. Interesting questions abound that call for research and academia–industry collaborations, such as which dynamic patterns in a film's content or plot development, such as peak-ebb-peak of narrative suspense or emotional intensity, will appeal the most to the audience? Does this vary by genre and audience segments? What characteristics of the conversation or discourse in a film would engage the audience the most? How would character development and chemistry among them influence the audience's engagement? Which music would enhance the artistic expressions of which scene? How much content overlap with or differentiation from the previous franchise would make a sequel most attractive?

6.1.2 AR/VR analytics

Advances of augmented and virtual realities (AR/VR) in the past decade have offered film production the potential to augment the traditional film-viewing experiences. AR/VR have received increased funding and been perceived by industry insiders as extremely important for film producers.[17] Interactivity brought by AR/VR has also emerged as a crucial element for consumer engagement and thus producer considerations (cf. Milgram and Kishino 1995). These technologies have also furnished creative tools to garner valuable data for producers, such as viewers' body motions, that may be leveraged to create more compelling content. Nonetheless, AR/VR applications and related academic research remain at infancy. Examples of extant applications include interactive films, such as Bandersnatch, during which audiences can actively decide how the plot unfolds, representing a current application. Other applications in the gaming industry, such as Quantic Dream, have also blurred the lines between films and games.

[17] VRVCA (2018). VR/AR global investment report and outlook 2018. Resource document. VRVCA. https://static1.squarespace.com/static/575e5cd62b8ddeb3fba63f79/t/5aa963d4e2c483ff9d60f3c9/1521050602030/VRVCA_Global+Investment+Report+2018_vF+%28EN%29.pdf.

Table 4 Important capabilities at the exploitation stage

Data acquisition	This pertains to textual and quantitative data regarding responses to the movie, both in terms of consumer and critical reviews (e.g., IMDb or Rotten Tomatoes), consumer discussions (e.g., on social platforms like Facebook or Reddit) or the movies general reception (e.g., sales data for utilized channels)
Data assimilation	This pertains to textual analytics (e.g., Humphreys and Wang 2017) such as sentiment analysis and topic modeling to gauge consumers' and critics' general reaction toward a film as well as common topics in their ratings and discussions; also, regression and descriptive analyses in order to gauge the performance in every channel and link them with previously discussed reactions
Data transformation	Data can be integrated into existing processes such as a film postmortem and add evidence and more fine-grained empirical insights into the team's learning from a particular movie's performance
Data exploitation	Value for consumers: By having their opinion about a film taken into account in a more structured way, following productions will target consumers' needs more directly, making content more compelling for the target audience
	Value for producers: With greater insights into consumer needs, producers can evaluate future movie productions more clearly in terms of attractiveness for the target audience and potential production value, which increases the odds for being profitable

Another related domain is the use of mobile devices in AR/VR, often termed the multi-device or multi-screen film-viewing experience. Nearly 81% of the US Internet users engage a second device, most favorably mobile, while watching TV.[18] Academic research is being developed to conceptualize multi-screen usage and understand related consumer behavior (e.g., Blake 2017; Jokela et al. 2015; Rooksby et al. 2015). Social media companies, such as Facebook, have poured significant resources into understanding why and how consumers use multiple devices during content consumption, particularly since most second-screen activities are related to social activities, such as texting friends or visiting social networks. Film producers can also adapt their content to transform the second-screen activities from distracting to engaging (Blake 2017; Hassoun 2014), potentially by arousing viewer conversations about the content and interactions among viewers. Moreover, enriched information or interactive quizzes about the content may be offered to permit dynamic audience reactions or participation in the story. Many intriguing questions remain in this domain. For instance, which kinds of augmented experiences are most welcome by the audience? How do they differ between the theatrical versus home entertainment experiences? Which parts of the film content are most suited for experience augmentation? How can producers generate supplemental content for secondary screens? How can producers leverage new forms of augmented experiences, such as bullet screens (where audiences input real-time comments to the displaying screen) that have become very popular in China and Japan, to engage the audience?

[18] IAB, and MARU/Matchbox (2017). The changing TV experience: 2017. Online report. Lab and maru/matchbox. https://www.iab.com/wp-content/uploads/2017/05/The-Changing-TV-Experience-2017.pdf.

6.1.3 Experience analytics

Central to all the above discussions on content analytics, AR/VR analytics, and AI-generated content, are the audience experience analytics. After all, understanding and optimizing the customer/audience experience is the bedrock of any successful content production. Beyond the conventional pre- and post-experiential measures, unobtrusive in-experience measures are becoming more accessible to content producers, such as facial expression, eye tracking, and seat movement tracked easily by film theaters. In addition, social experience, particularly among a larger number of consumers as in the case of theater experience, is not well explored by the marketing literature (with a few exceptions such as Ramanathan and McGill 2007; Ratner and Hamilton 2015; Kim et al. 2019a, b). Such analytics will help film producers better understand the differences between the theatrical and home experiences; and the interaction between film content and social context, offering theatrical films a potentially distinct competitive edge over SVODs. Other fruitful areas of exploration may pertain to how producers may better harness the moment-to-moment experience data and analytics to adapt film content? How various content formats, such as short films, impact consumer experience? How would different viewing devices, such as mobile, impact the content type or format preference by viewers (Chisholm et al. 2015)?

6.1.4 Talent analytics

Another exciting area for future research and applications may rest on leveraging analytics for talent management. Instead of relying on talent's and agents' intuition, or simple statistics, talent analytics, often involving talent's image analytics, bankability analytics, and social media analytics, can enlist much richer historical data and more sophisticated methodologies to analyze a talent's fit with specific film content, fit with other talents, profit potential, and long-term career path. These measures are crucial for guiding the entire film production process from greenlight and financial structuring. In addition, while human resource analytics have largely focused on individuals, newer development has highlighted the importance of human network analytics, where each talent is analyzed within his/her specific groups, teams, and larger organizational or industrial network (Packard et al. 2016). Future research should further explore, for instance, how to select talent for specific content, or for franchise films? How to find the optimal on-screen dyad with the best chemistry? How can we select talent with the most Oscar potential? Which talent may generate the most publicity or social media frenzy?

In summary, the immense potential to explore analytics in film production, thus offering production companies an essential competitive asset, and the dearth of academic research and applications thus far, call for increased collaborations between the industry and academia, as well as interdisciplinary research across, for example, finance, marketing, computer science, psychology, and linguistics. Producers also need to be open-minded and ready to adopt new data and new methods to facilitate profitable film production and to remain competitive in the rapidly changing digital media landscape.

6.2 Despite the promise offered by analytics, a number of impediments potentially curtail their wider applications and impact

Specifically, these impediments may arise from data restrictions, methodological challenges, talent limitations, affordability concerns, and organizational resistance. We will concisely describe each below.

6.2.1 Data

If the data needed to implement the desired analytic methods are not available, of poor quality, or difficult to make sense of, then a producer's capacity becomes restricted. Filmmaking is collaborative, but an individual producer, including one hired or financed by a major company is often at a remove from key audience and performance data. This problem is further highlighted for independent producers (Fuselier 2017). The key question remains—"How can you be data driven when someone else has the customer data?" For studios, the "someone else" may include the exhibitors and SVOD platforms. For independent producers, the "someone else" is "everyone else." Therefore, research is clearly required to accompany industry initiatives to deliver greater data transparency.[19] Issues of data access not only concern individual companies, but filmmaking cultures as a whole. For instance, the lack of harmonized accounting standards, poor information transparency, and shortage of adequate industry databases have been noted as the key reasons for the absence of expansion to the European audiovisual financing market (Debande 2018).

6.2.2 Methodology

Another challenge facing producers is a better understanding of the often complex analytic methods and then applying the resulting insights to decision-making. For example, applying artificial intelligence to audio/video content analysis poses challenges for producers.[20] In addition, the continuous, iterative nature of applying and improving analytics is sometimes deemed at odds with the established work practices of filmmaking, which are typically one-off productions. This speaks to the necessity of access to scale and continuous activities beyond the reach of most non-vertically integrated producers.

6.2.3 Analytics talent

Without establishing a solid understanding of the research objects and questions, the undifferentiated applications of analytics will inadvertently lead to unsatisfactory

[19] EFM Horizon (2019). Blockchain in Motion. https://www.efm-berlinale.de/en/horizon/programme/blockchain/blockchain.html#!/accordion1085974=accordion-item-start-module+accordion-item-start-module-1.

[20] Dixon (2019).Why recommendations aren't better, yet! NScreen Media. http://www.nscreenmedia.com/boost-metadata-better-video-recommendations/.

and potentially misleading insights for producers. Thus, combining expert knowledge, academic theory, and empirical findings is imperative to choose suitable methods and data sources in order to reach sensible conclusions to complement managerial intuitions (Hennig-Thurau and Houston 2018). Then successful implementations of analytics will require recruiting data science talent, who not only is already in great shortage on the global labor market but also incurs great costs to producers.

6.2.4 Affordability

Curating, storing, and analyzing big data with massive volume, variety, velocity (i.e., speed of data generation) and veracity (data uncertainty) generate cost concerns, especially for smaller production companies. As the industry's 2018 Nostradamus Report noted, "controlling the whole value chain from production to eyeballs has enormous value, allowing one to own the audience relationship as well as all the user data. This is why platforms like Netflix, Apple, Facebook and Snapchat now produce content of their own (and why most TV networks always have). Another strategy is for the content provider to build its own platforms, like Disney is doing." Such strategies to maximize the exploitation of leveraging analytics are apparently beyond the reach of most companies outside global entities.

6.2.5 Organizational resistance

It is not uncommon that creative personnel believe artistic decisions should not be informed by analytics. The concerns about AI-generated creative content and relegation of creative practice are reflected most deeply at the ideation stage. Some fear redundancy as a result of data-driven content generation. Others in computational creativity research have recognized the concern about the technical capacity and sociological hurdles for accepting the creative agency of software and proposed framings for its comprehension (Charnley et al. 2012). It is an area of ongoing debate in the film industry, where resistance to data-driven or data-informed approaches is both a legitimate live issue but also a topic of controversy promulgated in the trade press.[21] Simon and Schroeder point out that the tension between artistic value of filmmakers and the rationalization of practice through the use of big data (Napoli 2010) will be bridged in so far as producers believe the use of analytics is vital to profitability. These sentiments point to the need for more applied research and nuanced accounts of multiple approaches toward, and contexts for, the relationship between data and creative practice (Hennig-Thurau and Houston 2018).

By presenting Sects. 6.3 and 6.4, we will also discuss how the industry, academia, and regulatory bodies might work together to mitigate some of the impediments in Sect. 6.2, address the conceptual and methodological sparsity of research, and accelerate the adoption of analytics in film production.

[21] For example, Caranicas (2018). Artificial intelligence could 1 day determine which films get made. Variety. https://variety.com/2018/artisans/news/artificial-intelligence-hollywood-1202865540.

6.3 Advances in data and analytics will deepen and widen their applications in film production

Technological advances have been rapidly improving the access to data and analytic techniques. Below, we will highlight a few recent developments and the implications for film production.

6.3.1 Wider and collaborative access to data

Big data of massive volume, variety, velocity, and veracity continue to emerge. Cloud-based databases are replacing the analog databases and can thus be more easily shared both among producers and across the value chain, such as exhibitors, distributors, and producers. Although these entities are largely not sharing data as of now, there has been an urgent call for them to share data to accomplish a common goal. As aforementioned, examples in China have demonstrated successful implementations of cross-value chain datasets where financiers, producers, ticketing apps, and Internet platforms can be cross-owned, enabling both marketing efficiencies and higher level strategic decision-making (Roxborough 2016). Many open-source data, including those crowd-sourced or generated by users, such as audience reviews, social media comments, and kaggle competitions, can also be easily curated and analyzed even by producers with relatively low budgets for analytics.

Producers may leverage not only data curated by upstream and downstream channel members, as well as from open sources, but also new, often termed "dark," data generated in the content production process, such as production logs or test audiences. These dark data are often defined as the information assets that organizations collect, process, and store during regular business activities, but generally fail to use for other purposes. They are often recorded and stored for compliance purposes, only without being monetized via analytics to gain a competitive advantage. Producers should consider harnessing the power of such data and increasingly shift from intuition- to analytics-driven decision-making. The marketing academia is also standing at an unparalleled position to catalyze this process.

6.3.2 Accessible and collaborative analytics

Open-source software, such as R and Python, and prepackaged analytics software, such as those text, image, and video analytics tools developed by Amazon, Google, and Facebook, have dramatically reduced the cost and learning threshold for anyone, producers included, to leverage analytics in decision-making, especially when these decisions pertain to multi-million dollar investment in a film. Such open-source or pre-built software is also leveling the playing field for organizations of varied size and capability. Analytics are now commonly utilized to help both big and small companies improve their business processes; and help humans make more precise decisions, freeing humans to focus better on more critical and more creative tasks. Big data and analytics are here to stay, and like almost every other sector, the film industry will have to adapt to the opportunity they bring.

6.3.3 Interpretable and real-time analytics

Besides being accessible and collaborative, analytics are also becoming increasingly interpretable and faster for real-time decision-making. Improved data storytelling and visualization tools, and increased use of cloud-based data integration platforms permit a more unified approach to data and more producers to have the ability to tell relevant, accurate stories with the data. Text analytics, such as topic modeling, and image analytics, such as neural network with attention, also allow producers to interpret key characteristics in scripts, consumer reviews, and films that appeal to the audience.

Moreover, as more real-time data continue to emerge and real-time decisions need to be made, such as the real-time augmented content offered to the audience, real-time adaption of film content, or interactive films, content producers need to further harness real-time data and real-time analytics. This in-the-moment decision-making is particularly critical for managing resources, resolving issues and, ultimately, delighting customers. Two notable technologies are helping accomplish this. One is edge computing, which reduces the time involved in storing, retrieving, processing, and analyzing data stored on cloud or a large number of connected devices. The other is quantum computing, which despite being in its infancy, offers promise to process billions of data at once in merely minutes, thus giving companies the opportunity to make timely decisions to achieve desired outcomes.

The above discussions point to both the necessity and promise of analytics to film production. They also highlight the need for industry–academia collaboration in this long, yet fruitful path. Many interesting questions remain unanswered, such as what is the most effective format for various players in the value creation chain to share resources, particularly granular, real-time data and analytic capabilities? Which distinctive competitive advantages may be created in this process? And which technologies may be leveraged to curate data throughout the film production process? Which other production decisions can be improved by enlisting collaborative or real-time analytics?

6.4 Enhanced regulatory support is needed to propel accelerated adoption of analytics in film production

Accompanying the rapid advances in data and analytics, regulatory support from both the industrial governing bodies and government legislation will further catalyze the adoption of analytics as strategic routines in film production. We will here emphasize three core components: regulations on data transparency and fair access, on data security and privacy protection, and on pro-social analytics.

6.4.1 Data transparency and fair access

The earlier discussions have highlighted the status quo that not all entities involved in the film production value chain have comparable access to data needed for analytics to enhance film production. There is an increased call for enhanced access to data from redesigning the Blockchain with a unique footprint of every contract, to partly address the skew in balance of power that the SVOD platforms hold via granular behavioral data. They also call for improving transparency between distributors and other market actors to increase the visibility and timeliness of performance data. Klossa's (2019) report has further proposed that European Commission launch a one billion euro media data initiative to focus on the applications of data analytics, AI, and media Blockchain.[22] It also urges mobilization of all stakeholders to collaborate in this process. Academic research is also needed to facilitate the process, such as demonstrating the overall benefit of this collaborative approach, designing viable mechanisms to accomplish fair access to data and analytics, and gauging the impact of this approach to the industrial organization and competitive landscape.

6.4.2 Data security and consumer privacy protection

In the past decade, more than 5000 major data breaches have been reported, with the recent examples of Sony and Marriott. The average cost of a data breach is estimated to be around $4 million (Wedel and Kanan 2016). These scandals highlight the importance of cybersecurity. Globally, European Union (EU) has enacted the General Data Protection Regulation (GDPR) since May 2018. Several US states are also proposing their own data protection laws that provide certain GDPR-like consumer rights, with the first passing in June 2018 in California. The California Consumer Privacy Act (CCPA) is slated to become the most comprehensive data privacy law in the USA and will go into effect on January 1, 2020. Eleven other states have recently introduced similar legislation. The bills include their own versions of opt-out rights and are slightly different than the GDPR and CCPA, leading many to ask the US Congress to implement a national comprehensive data privacy legislation to reduce the complexities involved in navigating across each state's regulations. As massive amount of individual viewers' preference and behavioral data are being generated and curated by content providers, self-regulations on data security and consumer privacy protection also need to be established.

The marketing literature has a historical, and newly revived, interest in consumer privacy due to the rise of big data and analytics (Wedel and Kannan 2016). As different forms of marketing data emerge over time, this literature has examined privacy concerns that arise in many marketing contexts and data forms, such as marketing research like surveys (Mayer and White Jr. 1969; Acquisti et al. 2012), direct marketing via phones or e-mails (Kumar et al. 2014; Goh et al. 2015), offline retail sales (Schneider et al. 2018), subscription services and various customer relationship management (CRM) programs (Conitzer et al. 2012), online personalization services in computers and mobile devices (Chellappa and Shivendu 2010), online search and transactions (i.e., e-commerce) (Bart et al. 2005), online social networks (e.g., Adjerid et al. 2019), health care (Miller and Tucker 2018; Adjerid et al. 2016), and digital advertising (Tucker 2014; Gardete and Bart 2018). Nonetheless, research

[22] Klossa (2019).Toward European Media Sovereignty. https://ec.europa.eu/commission/sites/beta-political/files/guillaume_klossa_report_final.pdf.

related to the creative industries remains sparse. Thus, future research may consider many topics of interest, such as how content producers may garner the audience's data for decision-making while preserving privacy? How may film production value chain share data to accomplish common objectives without sacrificing data security? How can the creative industries design optimal privacy-friendly policies toward consumers, such as allowing them to control privacy settings (Adjerid et al. 2019)? How to design a B2B market that preserves privacy yet incentivizes competitors to participate in data and analytics knowledge sharing (Kalvenes and Basu 2006)? How do privacy regulations impact content producers' adoption of analytics and innovative motivations?

6.4.3 Pro-social analytics

While our discussions have centered on leveraging analytics to produce appealing and profitable film content, it is also important to leverage analytics responsibly and pro-socially. Many interesting questions await the academia and executives to address, such as how to mitigate the potential bias, such as gender and racial bias, introduced by analytics in ideation or casting (Zhao et al. 2019)? How to enlist analytics, such as social media analytics, to generate more pro-social entertainment content for the audience, especially for the younger audience? How to leverage analytics to moderate undesirable consumer behavior, such as binge viewing or preference for violent content, and promote consumer well-being (Gordon et al. 2015; Schweidel and Moe 2016)? How to manage information revelation to viewers when certain content is generated by artificial instead of human intelligence and creativity?

6.4.4 Summary

Built upon the theories of value creation and film production, this article synthesizes the state-of-the-art research on, and applications of, analytics across three primary stages of film production from a producer's perspective. It further illuminates the need for film producers to leverage analytics, the potential impediments that producers face, and promising avenues for future research and regulatory support, thus calling for academia–industry–regulation collaborations.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Abadi, M. K., Staiano, J., Cappelletti, A., Zancanaro, M., & Sebe, N. (2013). Multimodal engagement classification for affective cinema. In *5th conference on affective computing and intelligent interaction (ACII)* (pp. 411–416).

Acheson, K., & Maule, C. J. (1994). Understanding Hollywood's organization and continuing success. *Journal of Cultural Economics, 18*(4), 271–300.

Acquisti, A., Leslie, K., & Loewenstein, G. (2012). The impact of relative standards on the propensity to disclose. *Journal of Marketing Research, 49*(4), 160–174.
Adjerid, I., Acquisti, A., & Loewenstein, G. (2019). Choice architecture, framing, and cascaded privacy choices. *Management Science, 65*(6), 2267–2290.
Adjerid, I., Acquisti, A., Telang, R., Padman, R., & Adler-Milstein, J. (2016). The impact of privacy regulation and technology incentives: The case of health information exchanges. *Management Science, 62*(4), 1042–1063.
Aguiar, L., & Waldfogel, J. (2018). Netflix: global hegemon or facilitator of frictionless digital trade? *Journal of Cultural Economics, 42*(3), 419–445.
Anfer, O., & Wamba, S. F. (2019). Big data analytics and strategic marketing capabilities: impact on firm performance. In *World conference on information systems and technologies* (pp. 633–640). Cham: Springer.
Asur, S., & Huberman, B. A. (2010). Predicting the future with social media. In *Proceedings of the 2010 IEEE/WIC/ACM international conference on web intelligence and intelligent agent technology* (Vol. 1, pp. 492–499).
Barbosu, S. (2017) Big data on the big screen: Revealing latent coherence among movies and its effect on box office. *Academy of Management Proceedings*.
Barnett, S. B., & Cerf, M. (2017). A ticket for your thoughts: Method for predicting content recall and sales using neural similarity of moviegoers. *Journal of Consumer Research, 44*(1), 160–181.
Barney, J. B. (1997). *Gaining and sustaining competitive advantage*. Michigan: Addison-Wesley Publishing Company.
Bart, Y., Shankar, V., Sultan, F., & Urban, G. (2005). Are the drivers and role of online trust the same for all web sites and consumers? A large–scale exploratory empirical study. *Journal of Marketing, 69*(10), 133–152.
Bilasco, I. M., Lablack, A., Dahmane, A., & Danisman, T. (2015). Analysing user visual implicit feedback in enhanced TV scenarios. In L. Agapito, M. Bronstein, & C. Rother (Eds.), *Computer Vision–ECCV 2014 Workshops* (pp. 315–324)., Lecture notes in computer science Cham: Springer.
Blake, J. (2017). Second screen interaction in the cinema: Experimenting with transmedia narratives and commercializing user participation. *Journal of Audience and Reception Studies, 14*(2), 526–544.
Bozik, K., & Dimovski, V. (2019). Business intelligence and analytics for value creation: The role of absorptive capacity. *International Journal of Information Management, 46,* 93–103.
Brandes, U., Hoefer, M., & Pich, C. (2006). Affiliation dynamics with an application to movie–actor biographies. In *IEEE–VGTC symposium on visualization* (Ed.) (pp. 176–186).
Brookey, R. A., & Zhang, Z. (2018). How Hollywood applies industrial strategies to counter market uncertainty: The issue of financing and exhibition. In *Handbook of state aid for film* (pp. 135–150). Cham: Springer.
Bruneel, C., Guy, J. L., Haughton, D., Lemercier, N., McLaughlin, M. D., Mentzer, K., & Zhang, C. (2018). Movie analytics and the future of film finance. Are Oscars and box office revenue predictable? In Murschetz, P. C., Teichmann, R., & Karmasin, M. (Ed.), *Handbook of state aid for film* (pp. 551–578). Cham: Springer.
Cabral, L., & Natividad, G. (2016). Box-office demand: The importance of being #1. *The Journal of Industrial Economics, 64*(2), 277–294.
Campo, M., Hsieh, C. K., Nickens, M., Espinoza, J. J., Taliyan, A., Rieger, J., Ho, J., & Sherick, B. (2018). Analysis system for theatrical movie releases based on movie trailer deep video representation. 20th Century Fox. Research Paper. arXiv:1807.04465v1[cs.IR].
Carr, A. (2013). *Positive psychology: The science of happiness and human strengths*. London: Routledge.
Castendyk, O. (2018). Tax incentive schemes for film production: A pivotal tool of film policy? In *Handbook of state aid for film* (pp. 597– 615). Cham: Springer.
Cattani, G., & Ferriani, S. (2008). A core/periphery perspective on individual creative performance: Social networks and cinematic achievements in the Hollywood film industry. *Organization Science, 19*(6), 824–844.
Cattani, G., Ferriani, S., Mariani, M. M., & Mengoli, S. (2013). Tackling the "Galácticos" effect: Team familiarity and the performance of star–studded projects. *Industrial and Corporate Change, 22*(6), 1629–1662.
Chamorro-Premuzic, T. (2016). *Personality and individual differences*. New York: Wiley.
Charnley, J. W., Pease, A., & Colton, S. (2012). On the notion of framing in computational creativity. In *International conference on computational creativity* (pp. 77–81).

Chellappa, R. K., & Shivendu, S. (2010). Mechanism design for “free” but “no free disposal” services: The economics of personalization under privacy concerns. *Management Science, 56*(10), 1766–1780.

Chintagunta, P., Gopinath, S., & Venkataraman, S. (2010). The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Science, 29*(5), 944–957.

Chisholm, D. C., Fernández-Blanco, V., Ravid, S. A., & Walls, W. D. (2015). Economics of motion pictures: The state of the art. *Journal of Cultural Economics, 39*(1), 1–13.

Cohen, M. (2017). Big data and service operations. *Production and Operations Management, 27*(9), 1709–1723.

Conitzer, V., Curtis, R. T., & Wagman, L. (2012). Hide and seek: Costly consumer privacy in a market with repeat purchases. *Marketing Science, 31*(2), 277–292.

Corts, K. (2001). The strategic effects of vertical market structure: Common agency and divisionalization in the US motion picture industry. *Journal of Economics and Management Strategy, 10*(4), 509–528.

Dalton, J. T., & Leung, T. C. (2017). Strategic decision-making in Hollywood release gaps. *Journal of International Economics, 105,* 10–21.

Davenport, T. H., Harris, J. G., & Morison, R. (2010). *Analytics at work: Smarter decisions, better results*. Cambridge: Harvard Business Press.

Dawar, N., & Parker, P. (1994). Marketing universals: Consumers’ use of brand name, price, physical appearance, and retailer reputation as signals of product quality. *Journal of Marketing, 58*(2), 81–95.

De Vany, A., & Walls, W. D. (1996). Bose-Einstein dynamics and adaptive contracting in the motion picture industry. *The Economic Journal, 106*(439), 1493–1514.

De Vany, A., & Walls, D. (2004). Motion picture profit, the stable Paretian hypothesis, and the curse of the superstar. *Journal of Economic Dynamics and Control, 28*(6), 1035–1057.

Debande, O. (2018). Film finance: The role of private investors in the European film market. In P. C. Murschetz, R. Teichmann, & M. Karmasin (Eds.), *Handbook of state aid for film* (pp. 51–66). Cham: Springer.

Delen, D., Sharda, R., & Kumar, P. (2007). Movie forecast Guru: A Web–based DSS for Hollywood managers. *Decision Support Systems, 43*(4), 1151–1170.

Deng, Z., Navarathna, R., Carr, P., Mandt, S., Yue, Y., Matthews, I., & Mori, G. (2017). Factorized variational autoencoders for modeling audience reactions to movies. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2577–2586).

Duan, W., Gu, B., & Whinston, A. (2008). Do online reviews matter? An empirical investigation of panel data. *Decision Support Systems, 45*(4), 1007–1016.

Einav, L. (2007). Seasonality in the US motion picture industry. *The Rand Journal of Economics, 38*(1), 127–145.

Ekman, P., & Friesen, W. V. (1978). *Facial action coding system: A technique for the measurement of facial movement*. Palo Alto: Consulting Psychologists Press.

Elberse, A. (2007). The power of stars: Do star actors drive the success of movies? *Journal of marketing, 71*(4), 102–120.

Elberse, A., & Eliashberg, J. (2003). Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures. *Marketing Science, 22*(3), 329–354.

Eliashberg, J., Elberse, A., & Leenders, M. A. (2006). The motion picture industry: Critical issues in practice, current research, and new research directions. *Marketing science, 25*(6), 638–661.

Eliashberg, J., Hui, S. K., & Zhang, Z. J. (2007). From story line to box office: A new approach for green–lighting movie scripts. *Management Science, 53*(6), 881–893.

Eliashberg, J., Hui, S. K., & Zhang, Z. J. (2014). Assessing box office performance using movie scripts: A kernel–based approach. *IEEE Transactions on Knowledge and Data Engineering, 26*(11), 2639–2648.

Finch, J., Horan, C., & Reid, E. (2015). The performativity of sustainability: making a conduit a marketing device. *Journal of Marketing Management, 31*(1–2), 167–192.

Fontaine, G., Grece, C., & Pumares, J. M. (2018). *Online video sharing: Offerings, audiences, economic aspects*. Strasbourg: European Audiovisual Observatory.

Foutz, N. (2017). Entertainment marketing. *Foundations and Trends in Marketing, 10*(4), 15–333.

Foutz, N. Z., & Jank, W. (2010). Pre-release forecasting for motion pictures using functional shape analysis of virtual stock markets. *Marketing Science, 29*(3), 568–579.

Galvão, M., & Henriques, R. (2018). Forecasting movie box office profitability. *Journal of Information Systems Engineering & Management, 3*(3), 22.
Gardete, P. M., & Bart, Y. (2018). Tailored cheap talk: The effect of privacy policy on ad content and market outcomes. *Marketing Science, 37*(5), 733–752.
Gaustad, T. (2018). How financing shapes a film project: Applying organizational economics to a case study in Norway. In *Handbook of state aid for film* (pp. 135–150). Cham: Springer.
Ghiassi, M., Lio, D., & Moon, B. (2015). Pre-production forecasting of movie revenues with a dynamic artificial neural network. *Expert Systems with Applications, 42*(6), 3176–3193.
Goh, K., Hui, K., & Png, I. (2015). Privacy and marketing externalities: Evidence from do not call. *Management Science, 61*(12), 2982–3000.
Goldman, W. (1983). *Adventures in the screen trade. A personal view of Hollywood and screenwriting*. New York: Warner Books.
Gomez-Uribe, C. A., & Hunt, N. (2016). The Netflix recommender system: Algorithms, business value, and innovation. *ACM Transactions on Management Information Systems, 6*(4), 13.
Gopinath, S., Chintagunta, P., & Venkataraman, S. (2013). Blogs, advertising, and local–market movie box office performance. *Management Science, 59*(12), 2635–2654.
Gordon, R., Gurrieri, L., & Chapman, M. (2015). Broadening an understanding of problem gambling: the lifestyle consumption community of sports betting. *Journal of Business Research, 68*(10), 2164–2172.
Grantham, B. (2012). Film risk in US motion pictures. In M. Hjort (Ed.), *Film and risk* (pp. 197–209). Detroit: Wayne State University Press.
Hababou, M., Amrouche, N., & Jedidi, K. (2016). Measuring economic efficiency in the motion picture industry: A data envelopment analysis approach. *Customer Needs and Solutions, 3*(3–4), 144–158.
Hadida, A. L. (2003). Strategic assets, institutional factors and performance: An application of the resource based view and of new institutional economics to cinema projects in France and the United States (1988–1997). *Unpublished doctoral dissertation*. Jouy en Josas: HEC.
Hadida, A. L. (2009). Motion picture performance: A review and research agenda. *International Journal of Management Reviews, 11*(3), 297–335.
Hanjalic, A., & Li–Qun, X. (2005). Affective video content representation and modeling. *IEEE Transaction on Multimedia, 7*(1), 143–154.
Hassoun, D. (2014). Tracing attentions: toward an analysis of simultaneous media use. *Television and New Media, 15*(4), 271–288.
Haughton, D., McLaughlin, M. D., Mentzer, K., & Zhang, C. (2014). Movie analytics: Visualization of the co–starring network. In *2014 IEEE 4th symposium on large data analysis and visualization (LDAV)* (pp. 115–116).
Hennig-Thurau, T. (2004). Spielfilme als anlageobjekte. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung, 56*(2), 171–188.
Hennig-Thurau, T., & Houston, M. (2018). *Entertainment science*. Heidelberg: Springer.
Hennig-Thurau, T., Houston, M., & Heitjans, T. (2009). Conceptualizing and measuring the monetary value of brand extensions: The case of motion pictures. *Journal of Marketing, 73*(4), 167–183.
Hennig-Thurau, T., Houston, M. B., & Sridhar, S. (2006). Can good marketing carry a bad product? Evidence from the motion picture industry. *Marketing letters., 17*(3), 205–219.
Hennig–Thurau, T., Völckner, F., Clement, M., & Hofmann. J. (2013). An ingredient branding approach to determine the financial value of stars: The case of motion pictures. SSRN Working Paper. http://dx.doi.org/10.2139/ssrn.1763547.
Hennig-Thurau, T., Walsh, G., & Wruck, O. (2001). An investigation into the factors determining the success of service innovations: The case of motion pictures. *Academy of Marketing Science Review, 6*(1), 1–23.
Hennig-Thurau, T., Wiertz, C., & Feldhaus, F. (2015). Does Twitter matter? The impact of microblogging word of mouth on consumers' adoption of new movies. *Journal of the Academy of Marketing Science, 43*(3), 375–394.
Hennig-Thurau, T., & Wruck, O. (2000). Warum wir ins Kino gehen–Erfolgsfaktoren von Kinofilmen. *Marketing ZFP, 22*(3), 241–258.
Hofmann, J., Clement, M., Völckner, F., & Hennig-Thurau, T. (2017). Empirical generalizations on the impact of stars on the economic success of movies. *International Journal of Research in Marketing, 34*(2), 442–461.
Houston, M. B., Kupfer, A. K., Hennig-Thurau, T., & Spann, M. (2018). Pre-release consumer buzz. *Journal of the Academy of Marketing Science, 46*(2), 338–360.

Humphreys, A., & Wang, R. J. H. (2017). Automated text analysis for consumer research. *Journal of Consumer Research, 44*(6), 1274–1306.

Hunter, I. I. I., David, S., Smith, S., & Singh, S. (2016). Predicting box office from the screenplay: A text analytical approach. *Journal of Screenwriting, 7*(2), 135–154.

Hur, M., Kang, P., & Cho, S. (2016). Box-office forecasting based on sentiments of movie reviews and Independent subspace method. *Information Sciences, 372,* 608–624.

Jannach, D., Resnick, P., Tuzhilin, A., & Zanker, M. (2016). Recommender systems—beyond matrix completion. *Communications of the ACM, 59*(11), 94–102.

Jansen, C. (2005). The performance of German motion pictures, profits and subsidies: Some empirical evidence. *Journal of Cultural Economics, 29*(3), 191–212.

Jokela, T., Ojala, J., & Olsson, T. (2015). A diary study on combining multiple information devices in everyday activities and tasks. In *CHI '15 proceedings of the 33rd annual ACM conference* (pp. 3903–3912).

Kalvenes, J., & Basu, A. (2006). Design of robust business–to–business electronic marketplaces with guaranteed privacy. *Management Science, 52*(11), 1721–1736.

Kim, E., Aribarg A., & Foutz, N. Z. (2019). How shared consumption enhances experiences. *Working paper.*

Kim, T., Hong, J., & Kang, P. (2015). Box office forecasting using machine learning algorithms based on SNS data. *International Journal of Forecasting, 31*(2), 364–390.

Kim, H., Katerenchuk, D., Billet, D., Huan, J., Park, H., & Li, B. (2019). Understanding actors and evaluating personae with Gaussian embeddings. In *Association for the advancement of artificial intelligence*.

Kosterich, A. (2016). Reconfiguring the "Hits": The new portrait of television program success in an era of big data. *International Journal on Media Management, 18*(1), 43–58.

Kumar, V., Zhang, X., & Luo, A. (2014). Modeling Customer Opt–In and Opt-Out in a Permission-Based Marketing Context. *Journal of Marketing Research, 51*(8), 403–419.

Kupfer, A. K., Pähler vor der Holte, N., Kübler, R. V., & Hennig-Thurau, T. (2018). The role of the partner brand's social media power in brand alliances. *Journal of Marketing, 82*(3), 25–44.

Lash, M. T., & Zhao, K. (2016). Early predictions of movie success: the who, what, and when of profitability. *Journal of Management Information Systems, 33*(3), 874–903.

Lee, S. Y., Qiu, L., & Whinston, A. (2018). Sentiment manipulation in online platforms: An analysis of movie tweets. *Production and Operations Management, 27*(3), 393–416.

Lehmann, D. R., & Weinberg, C. B. (2000). Sales through sequential distribution channels: an application to movies and videos. *Journal of Marketing, 64*(3), 18–33.

Liu, Y. (2006). Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of Marketing, 70*(3), 74–89.

Liu, L., Dzyabura, D., & Mizik, N. (2018). Visual Listening: brand monitoring from consumer created images. Working paper. University of Colorado.

Liu, A., Liu, Y., & Mazumdar, T. (2014). Star power in the eye of the beholder: A study of the influence of stars in the movie industry. *Marketing Letters, 25*(4), 385–396.

Liu, X., Shi, S. W., Teixeira, T., & Wedel, M. (2018b). Video content marketing: The making of clips. *Journal of Marketing, 82*(4), 86–101.

Lobato, R. (2018). Rethinking international TV flows research in the age of Netflix. *Television and New Media, 19*(3), 241–256.

Lovallo, D., Clarke, C., & Camerer, C. (2012). Robust analogizing and the outside view: Two empirical tests of case–based decision making. *Strategic Management Journal, 33*(5), 496–512.

Manovich, L. (2018). Digital traces in context| 100 billion data rows per second: Media analytics in the early 21st century. *International Journal of Communication, 12,* 473–488.

Marshall, P., Dockendorff, M., & Ibáñez, S. (2013). A forecasting system for movie attendance. *Journal of Business Research, 66*(10), 1800–1806.

Mayer, C. S., & White, C. H., Jr. (1969). The law of privacy and marketing research. *Journal of Marketing, 33*(4), 1–4.

McKenzie, J. (2012). The economics of movies: A literature survey. *Journal of Economic Surveys, 26*(1), 42–70.

McKenzie, J. (2013). Predicting box office with and without markets: Do internet users know anything? *Information Economics and Policy, 25*(2), 70–80.

Milgram, P., & Kishino, F. (1995). A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems, 77*(12), 1321–1329.

Miller, D., & Shamsie, J. (1996). The resource–based view of the firm in two environments: The Hollywood film studios from 1936 to 1965. *Academy of Management Journal, 39*(3), 519–543.
Miller, D., & Shamsie, J. (2001). Learning across the life cycle: experimentation and performance among the Hollywood studio heads. *Strategic Management Journal, 22*(8), 725–745.
Miller, A. R., & Tucker, C. (2018). Privacy protection, personalized medicine, and genetic testing. *Management Science, 64*(10), 4648–4668.
Napoli, P. (2010). *Audience evolution: New technologies and the transformation of media audiences*. New York: Columbia University Press.
Napoli, P. M. (2013). The algorithm as institution: Toward a theoretical framework for automated media production and consumption. *McGannon Center Working Paper Series*, (pp. 26), Fordham University Donald McGannon Communication Research Centre.
Noam, E. M. (2018). *Managing media and digital organizations*. Cham: Palgrave Macmillan.
Opitz, C., & Hofmann, K. H. (2014). Adverse selection and moral hazard in equity partnerships: Evidence from Hollywood's slate financing agreements. *Journal of Economics and Management Strategy, 23*(4), 811–838.
Packard, G., Aribarg, A., Eliashberg, J., & Foutz, N. Z. (2016). The role of network embeddedness in film success. *International Journal of Research in Marketing, 33*(2), 328–342.
Palmgreen, P., & Rayburn, J. D. (1982). Gratifications sought and media exposure: An expectancy value model. *Communication Research, 9*(4), 561–580.
Pennock, D. M., Lawrence, S., Giles, C. L., & Nielsen, F. A. (2001). The real power of artificial markets. *Science, 291*(5506), 987–988.
Peukert, C. (2018). The next wave of digital technological change and the cultural industries. *Journal of Cultural Economics, 43,* 189–210.
Phillips, R. (2004). The global export of risk: Finance and the film business. *Competition and Change, 8*(2), 105–136.
PwC. (2018). *Global entertainment & media Outlook 2018–2022*. New York: PricewaterhouseCoopers.
Ramanathan, S., & McGill, A. (2007). Consuming with others: Social influences on moment–to–moment and retrospective evaluations of an experience. *Journal of Consumer Research, 34*(4), 506–524.
Ramdarshan Bold, R. M. (2016). The return of the social author: Negotiating authority and influence on Wattpad. *Convergence: The International Journal of Research into New Media Technologies, 24*(2), 117–136.
Ratner, R. K., & Hamilton, R. W. (2015). Inhibited from bowling alone. *Journal of Consumer Research, 42*(2), 266–283.
Ravid, S. A. (1999). Information, blockbusters, and stars: a study of the film industry. *Journal of Business, 72,* 463–492.
Ravid, S. A. (2013). What makes movies tick Ivory tower insights, studio views and research directions. In Burger–Helmchen, T. (Ed.), *The economics of creativity: Ideas, firms and markets* (pp. 60–339).
Ravid, S. A. (2018). The economics of film financing: an introduction. In P. C. Murschetz, R. Teichmann, & M. Karmasin (Eds.), *Handbook of state aid for film* (pp. 39–49). Cham: Springer.
Ravid, S. A., & Basuroy, S. (2004). Managerial objectives, the R–rating puzzle, and the production of violent films. *Journal of Business, 77,* 155–192.
Rhee, T. G., & Zulkernine, F. H. (2016). Predicting movie box office profitability: A neural network approach. In *15th IEEE international conference on machine learning and applications (ICMLA)* (pp. 665–670).
Rooksby, J., Smith, T. E., Morrison, A., Rost, M., & Chalmers, M. (2015). Configuring attention in the multiscreen living room. In *ECSCW 2015: Proceedings of the 14th European conference on computer supported cooperative work* (pp. 243–261).
Schneider, M. J., Jagpal, S., Gupta, S., Li, S., & Yu, Y. (2018). A flexible method for protecting marketing data: An application to point–of–sale data. *Marketing Science, 37*(1), 153–171.
Schweidel, D. A., & Moe, W. (2016). Binge watching and advertising. *Journal of Marketing, 80*(5), 1–19.
Sharma, R., Shanks, G., Seddon, P., & Reynolds, P. (2010). The impact of strategy and maturity on business analytics and firm performance: a review and research agenda. *Association for Information Systems.* https://aisel.aisnet.org/cgi/viewcontent.cgi?referer=&httpsredir=1&article=1056&context=acis2010.
Simon, F. M., & Schroeder, R. (2019). Big data goes to Hollywood: The emergence of big data as a tool in the American film industry. In J. Hunsinger, M. Allen, & L. Klastrup (Eds.), *Second international handbook of internet research*. Dordrecht: Springer.

Soleymani, M., Pantic, M., & Pun, T. (2012). Multimodal emotion recognition in response to videos. *IEEE Transactions on Affective Computing (TAC), 3*(2), 211–223.

Spann, M., & Skiera, B. (2003). Internet-based virtual stock markets for business forecasting. *Management Science, 49,* 1310–1326.

Tashman, P., Marano, V., & Babin, J. (2019). Firm-specific assets and the internationalization–performance relationship in the US movie studio industry. *International Business Review, 28*(4), 785–795.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal, 18*(7), 509–533.

Toubia, O., Iyengar, G., Bunnell, R., & Lemaire, A. (2019). Extracting features of entertainment products: A guided LDA approach informed by the psychology of media consumption. *Journal of Marketing Research, 56*(1), 18–36.

Townsend, N. (2018). Working title films and universal: The integration of a British production company into a Hollywood studio. *Journal of British Cinema and Television, 15*(2), 179–203.

Tucker, C. E. (2014). Social networks, personalized advertising, and privacy controls. *Journal of Marketing Research, 51*(10), 546–562.

Van Roy, B., & Yan, X. (2010). Manipulation robustness of collaborative filtering. *Management Science, 56*(11), 1911–1929.

Verhoef, P. C., Kooge, E., & Walk, N. (2016). *Creating value with big data analytics: Making smarter marketing decisions*. London: Routledge.

Von Rimscha, M. B. (2009). Managing risk in motion picture project development. *Journal of Media Business Studies, 6*(4), 75–101.

Walls, W. D. (2005a). Modelling heavy tails and skewness in film returns. *Applied Financial Economics, 15*(17), 1181–1188.

Walls, W. D. (2005b). Modeling movie success when ‘nobody knows anything’: Conditional stable-distribution analysis of film returns. *Journal of Cultural Economics, 29*(3), 177–190.

Walraven, J. (2018). Data science and entertainment production. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 2879–2879).

Wamba, S. F., Gunasekaran, A., Akter, S., Ren, S. J., Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research, 70*(1), 356–365.

Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data–rich environments. *Journal of Marketing, 80*(11), 97–121.

Yang, J., & Yecies, B. (2016). Mining Chinese social media UGC: A big data framework for analyzing Douban movie reviews. *Journal of Big Data, 3*(3), 1–23.

Yeap, J., Ignatius, J., & Ramayah, T. (2014). Determining consumers’ most preferred eWOM platform for movie reviews: A fuzzy analytic hierarchy process approach. *Computers in Human Behavior, 31,* 250–258.

Yeh, S. Y. (2013). When do original movies dominate sequels? The moderating effects of film types and product knowledge. *Asia Pacific Management Review, 18*(3), 239–255.

Young, S. M., Gong, J. J., & Van der Stede, W. A. (2008). The business of making movies. *Strategic Finance, 89*(8), 26–32.

Zhao, J., Wang, T., Yatskar, M., Cotterell, R., Ordonez, V., & Chang, K. (2019). Gender bias in contextualized word embeddings. In *North American Chapter of the Association for Computational Linguistics*.

Zufryden, F. (1996). Linking advertising to box office performance of new film releases. *Journal of Advertising Research, 36*(4), 29–41.

Affiliations

Ronny Behrens[1] · Natasha Zhang Foutz[2] · Michael Franklin[3] · Jannis Funk[4] · Fernanda Gutierrez-Navratil[5] · Julian Hofmann[6] · Ulrike Leibfried[7]

Natasha Zhang Foutz
nfoutz@virginia.edu

Michael Franklin
m.franklin@gold.ac.uk

Jannis Funk
jannis.funk@filmuniversitaet.de

Fernanda Gutierrez-Navratil
luisafernanda.gutierrez@unavarra.es

Julian Hofmann
jhofmann@em-normandie.fr

Ulrike Leibfried
ulrike.leibfried@googlemail.com

1 Chair for Marketing and Media, University of Münster, Am Stadtgraben 13-15, 48143 Münster, Germany

2 McIntire School of Commerce, University of Virginia, 340 Rouss and Robertson Hall, Charlottesville, VA 22904, USA

3 Institute for Creative and Cultural Entrepreneurship (ICCE), Goldsmiths, University of London, New Cross, London SE14 6NW, UK

4 Film University Babelsberg KONRAD WOLF, Marlene-Dietrich-Allee 11, 14482 Potsdam, Germany

5 Universidad Pública de Navarra, Campus de Arrosadia, 31006 Pamplona, Spain

6 EM Normandie Business School, Métis Lab, 30 rue de Richelieu, 76087 Le Havre Cedex, France

7 Area 46 Development GmbH, Winsstraße 46, Berlin, Germany